January 1989 DTIC FILE COPY UILU-ENG-89-2203

(2)

COORDINATED SCIENCE LABORATORY
College of Engineering

SIGNAL DETECTION IN FRACTIONAL GAUSSIAN NOISE AND AN RKHS APPROACH TO ROBUST DETECTION AND ESTIMATION

Richard James Barton

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS None
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited
4. PERFORMING ORGANIZATION REPORT NUMBER(S) UILU-ENG-89-2203	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Coordinated Science Lab University of Illinois 6b. OFFICE SYMBOL (If applicable) N/A	7a. NAME OF MONITORING ORGANIZATION Office of Naval Research National Science Foundation
6c. ADDRESS (City, State, and ZIP Code) 1101 W. Springfield Ave. Urbana, IL 61801	7b. ADDRESS (City, State, and ZIP Code) 800 N. Quincy, Arlington, VA 22217 1800 G. St., Washington, D.C. 20552
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Office of Naval Research/National Science Found. 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-81-K-0014, N00014-89-J-1321 ECS-85-12314
8c. ADDRESS (City, State, and ZIP Code) 800 N. Quincy, Arlington, VA 22217 1800 G. St., Washington, D.C. 20552	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. | PROJECT NO. | TASK NO. | WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Signal Detection in Fractional Gaussian Noise and an RKHS Approach to Robust Detection and Estimation

12. PERSONAL AUTHOR(S)
Barton, Richard James

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM ______ TO ______	February 1989	146

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	signal detection, fractal noise, reproducing kernel Hilbert spaces, robust detection ~~and estimation~~, Radio communication. JES

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This report is divided into two parts. In the first part, the problem of signal detection in fractional Gaussian noise is considered. To facilitate the study of this problem, several results related to the reproducing kernel Hilbert space of fractional Brownian motion are presented. In particular, this reproducing kernel Hilbert space is characterized completely, and an alternative characterization for the restriction of this class of functions to a compact interval, $[O,T]$ is given. Infinite-interval whitening filters for fractional Brownian motion are also developed. Application of these results to the signal detection problem yields necessary and sufficient conditions for a deterministic or stochastic signal to produce a nonsingular shift when embedded in additive fractional Gaussian noise. Also, a formula for the likelihood ratio corresponding to any deterministic nonsingular shift is developed. Finally, some results concerning detector performance in the presence of additive fractional Gaussian noise are presented.

(continued)

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL

DD Form 1473, JUN 86 — Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

UNCLASSIFIED

19. Abstract (continued)

In the second part of the report, the application of reproducing kernel Hilbert space theory to the problems of robust detection and estimation is investigated. It is shown that this approach provides a general and unified framework in which to analyze the problems of L^2 estimation, matched filtering, and quadratic detection in the presence of uncertainties regarding the second-order structure of the random processes involved. Minimax robust solutions to these problems are characterized completely, and some results concerning existence of robust solutions are presented.

SIGNAL DETECTION IN FRACTIONAL GAUSSIAN NOISE
AND
AN RKHS APPROACH TO ROBUST DETECTION AND ESTIMATION

BY

RICHARD JAMES BARTON

A.B., University of Illinois, 1976
M.S., University of Illinois, 1984

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1989

Urbana, Illinois

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

THE GRADUATE COLLEGE

DECEMBER 1988

WE HEREBY RECOMMEND THAT THE THESIS BY

RICHARD JAMES BARTON

ENTITLED SIGNAL DETECTION IN FRACTIONAL GAUSSIAN NOISE AND AN RKHS APPROACH TO ROBUST DETECTION AND ESTIMATION

BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

Director of Thesis Research

N. Narayana Rao

Head of Department

Committee on Final Examination†

K. S. Arun — Chairperson

Andrew R. Barron

† Required for doctor's degree but not for master's.

0-517

ABSTRACT

This thesis is divided into two parts. In the first part, the problem of signal detection in fractional Gaussian noise is considered. To facilitate the study of this problem, several results related to the reproducing kernel Hilbert space of fractional Brownian motion are presented. In particular, this reproducing kernel Hilbert space is characterized completely, and an alternative characterization for the restriction of this class of functions to a compact interval $[0,T]$ is given. Infinite-interval whitening filters for fractional Brownian motion are also developed. Application of these results to the signal detection problem yields necessary and sufficient conditions for a deterministic or stochastic signal to produce a nonsingular shift when embedded in additive fractional Gaussian noise. Also, a formula for the likelihood ratio corresponding to any deterministic nonsingular shift is developed. Finally, some results concerning detector performance in the presence of additive fractional Gaussian noise are presented.

In the second part of the thesis, the application of reproducing kernel Hilbert space theory to the problems of robust detection and estimation is investigated. It is shown that this approach provides a general and unified framework in which to analyze the problems of L^2 estimation, matched filtering, and quadratic detection in the presence of uncertainties regarding the second-order structure of the random processes involved. Minimax robust solutions to these problems are characterized completely, and some results concerning existence of robust solutions are presented.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor H. V. Poor, for his guidance and encouragement, without which this work would not have been possible. I would also like to thank the other members of my thesis committee for many helpful and stimulating discussions, both in and out of the classroom.

Thanks are due to many of my fellow graduate students on the third floor of the Coordinated Science Lab for tolerance, advice, friendship, and esprit de corps during the Bears' games. This holds in particular for my officemates Behnaam Aazhang, Kapil Chawla, Arvind Krishna, Galen Sasaki, and Venu Veeravalli.

Last but not least, I would like to thank my wife Shelley, for everything.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In this thesis, we consider a variety of problems in the area of statistical signal processing. The work is divided into two relatively unrelated parts, the common factor being the application of reproducing kernel Hilbert space (RKHS) theory to the various problems considered. Since each part of the thesis is essentially self-contained, each comprising a separate chapter, we present here only some brief introductory remarks and an overview of the thesis. Each of the two major chapters includes a more complete introduction to the material presented.

In the first part of the thesis, we consider the problem of communication in the presence of noise displaying strong long-term dependence. We are motivated to consider this problem by the prevalence of natural phenomena that exhibit behavior indicative of such dependence. Among the many physical processes exhibiting such behavior are river flows, frequency fluctuations in oscillators, current fluctuations in semiconductors, and errors on communications channels. Given the ubiquitous nature of phenomena displaying long-term dependence, it seems desirable to consider the problem of communication in the presence of strongly dependent noise. We consider one aspect of this problem, namely, signal detection in the presence of additive, strongly dependent noise, and we adopt as our noise model the class of random processes known as fractional Brownian motions and the associated derivative processes referred to as fractional Gaussian noises.

In the second part of the thesis, we consider the problems of robust L^2 estimation, matched filtering, and quadratic detection in the presence of uncertainty regarding the statistical structure of the random processes involved. In recent years, the game-theoretic minimax approach to designing robust detection and estimation procedures has been studied by many authors. We also employ a minimax strategy, but we formulate and analyze the robust detection and estimation problems in an RKHS context. This approach provides a general and unified framework in which to analyze these problems and also clearly reveals the underlying similarities of the problems.

The thesis is organized as follows. In Chapter 2, we give a brief review of the relevant RKHS theory. In Chapter 3, we study the problem of signal detection in fractional Gaussian noise, and in Chapter 4, we consider the robust detection and estimation problems. In Chapter 5, we present some concluding remarks.

CHAPTER 2

REVIEW OF REPRODUCING KERNEL HILBERT SPACE RESULTS

In subsequent chapters of this thesis, we make frequent use of some basic results from the theory of reproducing kernel Hilbert spaces. We give below a brief review of these results. For a thorough introduction to the subject, see, for example, [74].

Let I be any index set. The term *covariance function* refers to any symmetric, nonnegative-definite function $K: I^2 \to \mathbb{C}$. Associated with any such covariance function, there is a unique Hilbert space $\boldsymbol{H}(K)$ of functions defined on I such that, for all $t \in I$ and all $f \in \boldsymbol{H}(K)$,

$$K(\cdot,t) \in \boldsymbol{H}(K),$$

and

$$f(t) = \langle f, K(\cdot,t) \rangle_{\boldsymbol{H}(K)}.$$

This function space is called the *reproducing kernel Hilbert space*, or RKHS, with *reproducing kernel K*, and it is well-known (see [1], §I.2) that $\boldsymbol{H}(K)$ consists of functions $f: I \to \mathbb{C}$ of the form

$$f(\cdot) = \sum_{i=1}^{N} \phi_i K(\cdot,t_i), \qquad \{\phi_i\}_{i=1}^{N} \subset \mathbb{C},\ \{t_i\}_{i=1}^{N} \subseteq I,$$

and their limits under the norm

$$\|f\|^2_{\boldsymbol{H}(K)} = \sum_{i,j=1}^{N} \phi_i \overline{\phi}_j K(t_j,t_i).$$

Note that norm convergence in $\boldsymbol{H}(K)$ implies pointwise convergence on I since, for all $t \in I$

and all $f, g \in \mathbf{H}(K)$,

$$|f(t) - g(t)| = |\langle (f - g), K(\cdot,t)\rangle_{\mathbf{H}(K)}| \\ \leq \|f - g\|_{\mathbf{H}(K)} K(t,t).$$

Given any two RKHSs $\mathbf{H}(K_0)$ and $\mathbf{H}(K_1)$ defined on the index set I, the *direct product space* $\mathbf{H}(K_0) \otimes \mathbf{H}(K_1)$ consists of functions $g : I^2 \to \mathbb{C}$ of the form

$$g(\cdot,*) = \sum_{i,j=1}^{N} \gamma_{ij} K_0(\cdot,t_i) K_1(*,t_j), \qquad \{\gamma_{ij}\}_{i,j=1}^{N} \subset \mathbb{C},\ \{t_i\}_{i=1}^{N} \subseteq I,$$

and their limits under the norm

$$\|g\|^2_{\mathbf{H}(K_0) \otimes \mathbf{H}(K_1)} = \sum_{i,j=1}^{N} \sum_{k,l=1}^{N} \gamma_{ij} \overline{\gamma_{kl}} K_0(t_k,t_i) K_1(t_l,t_j).$$

It follows (see [1], §I.8) that $\mathbf{H}(K_0) \otimes \mathbf{H}(K_1)$ is itself an RKHS with reproducing kernel

$$K((t_1,\tau_1);\ (t_2,\tau_2)) = K_0(t_1,t_2) K_1(\tau_1,\tau_2), \qquad (t_i,\tau_i) \in I^2.$$

If $\mathbf{X} \triangleq \{X(t);\ t \in I\}$ is a stochastic process with covariance function $K_{\mathbf{X}}$ and mean function $m \in \mathbf{H}(K_{\mathbf{X}})$, then $\mathbf{H}(\mathbf{X}) \triangleq \mathbf{H}(K_{\mathbf{X}})$ is congruent (i.e., isometrically isomorphic) to the Hilbert space $L^2(\mathbf{X})$ spanned by the random variables $\{X(t),\ t \in I\}$ (see [48], §2). For any $g \in \mathbf{H}(\mathbf{X})$, the corresponding element in $L^2(\mathbf{X})$ is usually denoted by $\langle \mathbf{X}, g\rangle_{\mathbf{H}(\mathbf{X})}$ and is characterized by the property that

$$g(t) = \mathrm{Cov}\left\{X(t), \langle \mathbf{X}, g\rangle_{\mathbf{H}(\mathbf{X})}\right\}, \qquad \forall\ t \in I. \tag{2.1.1}$$

It follows that, for all $t \in I$ and $g, h \in \mathbf{H}(K_{\mathbf{X}})$,

$$\langle \mathbf{X}, K_{\mathbf{X}}(\cdot,t)\rangle_{\mathbf{H}(\mathbf{X})} = X(t),$$

$$\mathrm{E}\left\{\langle \mathrm{X}, g\rangle_{\boldsymbol{H}(\mathrm{X})}\right\} = \langle m, g\rangle_{\boldsymbol{H}(\mathrm{X})}, \tag{2.1.2}$$

and

$$\mathrm{Cov}\left\{\langle \mathrm{X}, g\rangle_{\boldsymbol{H}(\mathrm{X})}, \langle \mathrm{X}, f\rangle_{\boldsymbol{H}(\mathrm{X})}\right\} = \langle h, g\rangle_{\boldsymbol{H}(\mathrm{X})}. \tag{2.1.3}$$

For our purposes, we will often take I itself to be a Hilbert space $\boldsymbol{H}_0$. For example, $\boldsymbol{H}_0$ might be $L^2(\boldsymbol{R})$, the space of real-valued functions that are square-integrable with respect to Lebesgue measure on $\boldsymbol{R}$. In this case, the observations are regarded as a *generalized random process*; that is, X is regarded as a linear operator mapping $\boldsymbol{H}_0$ into the space of square-integrable random variables on some probability space. The space $L^2(\mathrm{X})$ then consists of mean-square limits of random variables of the form $\mathrm{X}(f)$, $f \in \boldsymbol{H}_0$, and a covariance function is a bilinear form $K: \boldsymbol{H}_0^2 \to \boldsymbol{R}$. In fact, we will assume that all covariance functions are bounded bilinear forms so that they are generated by *covariance operators* on $\boldsymbol{H}_0$; that is, for all $f, g \in \boldsymbol{H}_0$,

$$K(f,g) = \mathrm{E}\{X(f)X(g)\} = \langle f, \mathrm{K}g\rangle_{\boldsymbol{H}_0},$$

where $\mathrm{K}: \boldsymbol{H}_0 \to \boldsymbol{H}_0$ is bounded, linear, self-adjoint, and positive (see [14], Lemma X.2.2). The RKHS corresponding to such a covariance function (denoted by $\boldsymbol{H}(\mathrm{K})$ as well as $\boldsymbol{H}(K)$) consists of bounded linear functionals $s: \boldsymbol{H}_0 \to \boldsymbol{R}$ of the form

$$s(\cdot) = \langle \cdot, \mathrm{K}g\rangle_{\boldsymbol{H}_0}, \qquad g \in \boldsymbol{H}_0,$$

and their limits under the norm

$$\|s\|^2_{\boldsymbol{H}(K)} = \langle g, \mathrm{K}g\rangle_{\boldsymbol{H}_0}.$$

Equivalently (see [48], §9), $\boldsymbol{H}(K)$ consists of functionals s of the form

$$s(f) = \langle \mathrm{K}^{1/2} f, S \rangle_{\boldsymbol{H}_0} = \langle f, \mathrm{K}^{1/2} S \rangle_{\boldsymbol{H}_0}, \qquad f \in \boldsymbol{H}_0,$$

where $\mathrm{K}^{1/2}$ is the square root of the operator K and $S \in \boldsymbol{H}_0$ is contained in the closure of the range of $\mathrm{K}^{1/2}$. The norm of s in $\boldsymbol{H}(K)$ is then

$$\|s\|^2_{\boldsymbol{H}(K)} = \langle S, S \rangle_{\boldsymbol{H}_0}.$$

Of course, appealing to the congruence between $\boldsymbol{H}_0$ and its dual $\boldsymbol{H}_0^*$, we could just as well say that $\boldsymbol{H}(K)$ consists of functions $s \in \boldsymbol{H}_0$ of the form $s = \mathrm{K}^{1/2} S$, where S is contained in the closure of the range of $\mathrm{K}^{1/2}$. While this is not technically correct (since $\boldsymbol{H}(K) \subset \boldsymbol{H}_0^*$), it is a useful way of thinking, and we will often blur the distinction between functions and functionals by writing statements such as

$$s(f) = \langle f, s \rangle_{\boldsymbol{H}_0}, \qquad f \in \boldsymbol{H}_0.$$

The following two results are useful for characterizing RKHSs and finding canonical representations of stochastic processes. Proofs can be found, for example, in [20].

Theorem 2.1.1: Let I be any index set and let K be a covariance function defined on I. Suppose there exists a measure space $(\Lambda, \boldsymbol{B}, \nu)$ and a set of functions $\{f_t;\ t \in I\} \subseteq L^2(\nu)$ such that

$$K(s,t) = \int_\Lambda f_s(\lambda) \overline{f_t(\lambda)}\, d\nu(\lambda), \qquad \forall\ (s,t) \in I^2.$$

Then $\boldsymbol{H}(K)$ consists of functions of the form

$$g(t) = \int_\Lambda f_t(\lambda) \overline{\tilde{g}(\lambda)}\, d\nu(\lambda), \qquad t \in I,$$

where $\tilde{g} \in \operatorname{span}\{f_t;\ t \in I\}$. Further, for all $g, h \in \boldsymbol{H}(K)$,

$$\langle g, h\rangle_{\boldsymbol{H}(K)} = \int_{\Lambda} \tilde{h}(\lambda)\overline{\tilde{g}(\lambda)}\,d\nu(\lambda).$$

Theorem 2.1.2: If $\mathbf{X} \triangleq \{X(t);\ t \in I\}$ is a zero-mean stochastic process with covariance function $K_{\mathbf{X}}$ such that, as above,

$$K_{\mathbf{X}}(s,t) = \int_{\Lambda} f_s(\lambda)\overline{f_t(\lambda)}\,d\nu(\lambda), \qquad \forall\ (s,t) \in I^2,$$

then there exists an orthogonal process $\mathbf{Z_X}$ on Λ with associated measure ν such that

$$X(t) = \int_{\Lambda} f_t(\lambda)\,dZ_{\mathbf{X}}(\lambda), \qquad \forall\ t \in I,$$

and $L^2(\mathbf{X}) = L^2(\mathbf{Z_X})$ if and only if $\{f_t;\ t \in I\}$ spans $L^2(\nu)$. Further, for all $g \in \boldsymbol{H}(\mathbf{X})$,

$$\langle \mathbf{X}, g\rangle_{\boldsymbol{H}(\mathbf{X})} = \int_{\Lambda} \tilde{g}(\lambda)\,dZ_{\mathbf{X}}(\lambda),$$

where $\tilde{g}$ is given by Theorem 2.1.1.

The final two results of this section relate RKHS theory to hypothesis testing. Theorem 2.1.3 is well-known, and excellent discussions of it can be found in [21] and [22]. Theorem 2.1.4 is due to Oodaira [45].

Theorem 2.1.3: Suppose that I is a separable metric space and $\mathbf{X} \triangleq \{X(t);\ t \in I\}$ is Gaussian with continuous covariance function $K_{\mathbf{X}}$. The hypothesis testing problem:

H_0: X has mean zero

versus

H_1: X has mean function m, is nonsingular if and only if $m \in \boldsymbol{H}(\mathrm{X})$, in which case, the likelihood ratio is given by

$$L(\mathrm{X}) = \exp\left[\langle \mathrm{X}, m \rangle_{\boldsymbol{H}(\mathrm{X})} - \tfrac{1}{2}\langle m, m \rangle_{\boldsymbol{H}(\mathrm{X})}\right]$$

Theorem 2.1.4: Let I be a separable metric space and let $\mathrm{X} \triangleq \{X(t);\ t \in I\}$ be a Gaussian process. Consider the following hypothesis testing problem:

H_0: X has mean zero and continuous covariance function K_0

versus

H_1: X has mean function m_1 and continuous covariance function K_1.

This problem is nonsingular if and only if the following three conditions are satisfied.

(i) $(K_1 - K_0) \in \boldsymbol{H}(K_0) \otimes \boldsymbol{H}(K_0)$.

(ii) $\boldsymbol{H}(K_0) = \boldsymbol{H}(K_1)$ (in the set theoretic sense), or equivalently, there exist constants $0 < c < C < \infty$ such that $(CK_0 - K_1)$ and $(K_1 - cK_0)$ are both nonnegative definite. (This is often abbreviated $cK_0 \ll K_1 \ll CK_0$.)

(iii) $m_1 \in \boldsymbol{H}(K_0)$.

CHAPTER 3

SIGNAL DETECTION IN FRACTIONAL GAUSSIAN NOISE

3.1. Introduction

In statistical signal processing applications, the lack of independence between observations has traditionally been handled by modeling the data as an ARMA process with relatively few parameters. Unfortunately, there are many physical processes that exhibit strong, positive, long-term correlations, which are not well-modeled by such ARMA processes. Such long-term dependence is very often observed, for example, in geophysical data, where it takes the form of long periods of high or low values (see [18] and the references cited therein). Similarly, errors on communications channels, "... appear to be grouped in bursts, which are in turn grouped in bursts, etc." [33]. This tendency for low or high values to be followed by other low or high values is often referred to as the Joseph effect. It is indicative of a process possessing a long memory and is perhaps best explained in terms of the spectral behavior of such a process. In particular, if the observed process is stationary, and the correlations between observations are positive and fall off so slowly that the covariance function is not integrable, then the spectral density of the process will be unbounded at the origin. The predominance of low-frequency power is the cause of the Joseph effect.

An important class of physical processes exhibiting strong long-term dependence are those with *1/f-type* spectral behavior; that is, spectral densities approximately proportional to f^{1-2H}, where f represents frequency and H is a constant in the range $\frac{1}{2} < H < \frac{3}{2}$. This type of spectral behavior is observed in a great many different phenomena, including, for example, the frequency fluctuations in electrical oscillators, the current fluctuations in metal films and semiconductor devices [76], and the loudness fluctuations in speech and music [72]. Many of these *1/f-noises*, as they are frequently called, appear to be relatively stationary and Gaussian; but for values of $H \geq 1$, it is not clear how such a spectrum should be interpreted (since no stationary, L^2 process could possess such a spectrum). Much discussion has been devoted to this problem (see, for example, [26], [34], and [46]), but it is far from being resolved. For values of H in the range $\frac{1}{2} < H < 1$, however, such processes can be modeled as generalized Gaussian processes or as the (stationary) increments of nonstationary Gaussian processes. This is, of course, analogous to the relationship between white Gaussian noise and ordinary Brownian motion. In this thesis, we restrict our attention to values of H in this range.

An early attempt at modeling *1/f*-noise was made by Barnes and Allan [5], who proposed modeling the phase noise in oscillators as a fractional integral of white noise. (See [44] for a discussion of fractional integration and differentiation.) The corresponding *1/f*-type frequency noise would then be modeled by the increments of the phase-noise process. The particular fractional integral that Barnes and Allan proposed was

$$\frac{1}{\Gamma(H+\frac{1}{2})}\int_0^t (t-\tau)^{H-\frac{1}{2}} dB(\tau), \qquad t \geq 0,$$

where $\{B(t);\, t \geq 0\}$ is a Brownian motion. Unfortunately, a process defined in this way does not have stationary increments. A later refinement of this model is the *fractional Brownian motion* (FBM) process introduced by Mandelbrot and Van Ness [36]. This process, which is discussed in more detail in the sequel, has stationary increments that exhibit *1/f*-type spectral behavior. In fact, in a certain sense, FBM has a stationary derivative, called *fractional Gaussian noise* (FGN), with spectral density equal to f^{1-2H}, $\frac{1}{2} < H < 1$. In addition to being stationary, the increments of FBM are also *self-similar*; that is, for each $a > 0$ and $t_0 \in \mathbb{R}$,

$$\{B_H(t_0+a\tau) - B_H(t_0);\ \tau \in \mathbb{R}\} \stackrel{d}{=} a^H \{B_H(\tau);\ \tau \in \mathbb{R}\},$$

where $\mathbf{B}_H$ is an FBM and "$\stackrel{d}{=}$" denotes equality of finite-dimensional distributions. Processes with self-similar and stationary increments have received much attention in recent years since they possess many interesting properties. In particular, the sample paths of such a process are *fractals* as defined in [35]. (For discussions of the various properties of these processes, see, for example, [27], [28], [42], [43], [60], [61], [62], [63], [64], and [71].) Because of its simplicity and its many interesting properties, FBM has become a popular statistical model. In addition to being the preeminent model for long-term dependence ([8], [18], [19], [29], and [37]), it is finding increasing application as a model for image texture ([32], [49], and [50]).

In this chapter, we consider the problem of detecting signals in the presence of additive FGN. We begin, in Section 3.2, with a discussion of the properties of FBM, including a rigorous treatment of FGN. In Section 3.3, we study the reproducing kernel Hilbert space of FBM. In Section 3.4, we discuss the problem of detecting deterministic signals in FGN, and in Section 3.5, we consider nondeterministic signals. In Section 3.6, we investigate some aspects of detector performance on FGN channels. Section 3.7 contains some concluding remarks.

3.2. Fractional Brownian Motion and Fractional Gaussian Noise

The class of *fractional Brownian motions*, or FBMs, was introduced by Mandelbrot and Van Ness in [36]. In this section we will define this class of processes and discuss some of their more interesting properties. A more complete development is given in [36].

For purposes of this paper, we will use the following (slightly specialized) version of the definition given in [36]. Motivation for this definition will be provided below. Let $\mathbf{B} \triangleq \{B(t);\ t \in \mathbb{R}\}$ be a standard Brownian motion and let $\frac{1}{2} < H < 1$. The fractional Brownian motion process $\mathbf{B}_H \triangleq \{B_H(t);\ t \in \mathbb{R}\}$ is defined by

$$B_H(t) \triangleq \frac{1}{\Gamma(H+\frac{1}{2})}\left[\int_{-\infty}^{0}\left(|t-\tau|^{H-\frac{1}{2}} - |\tau|^{H-\frac{1}{2}}\right)dB(\tau) + \int_{0}^{t}|t-\tau|^{H-\frac{1}{2}}dB(\tau)\right], \qquad t \in \mathbb{R}. \tag{3.2.1}$$

(Where, for $t < 0$, the notation " $\int_0^t$ " should be interpreted as " $-\int_t^0$ ".) Clearly, $\mathbf{B}_H$ is a zero-mean Gaussian process and $B_H(0) = 0$. Notice that if we extend our definition to include $H = \frac{1}{2}$, we get

$$B_{\frac{1}{2}}(t) = B(t), \qquad t \in \mathbb{R}.$$

In this sense, FBM can be regarded as a generalization of the familiar Brownian motion process. It is a generalization that is particularly useful for applications, as we shall see in the sequel.

The covariance function of $\mathbf{B}_H$ is given, after some analysis, by

$$K_{\mathbf{B}_H}(s,t) = \frac{V_H}{2}\left\{|s|^{2H} + |t|^{2H} - |t-s|^{2H}\right\}, \qquad s,t \in \mathbb{R}, \tag{3.2.2}$$

where

$$V_H \triangleq \mathrm{Var}[B_H(1)] = \frac{-\Gamma(2-2H)\cos(\pi H)}{\pi H(2H-1)}. \tag{3.2.3}$$

It follows from this covariance structure that the increments of $\mathbf{B}_H$ are stationary and *self-similar*; that is, for each $a > 0$ and $t_0 \in \mathbb{R}$,

$$\{B_H(t_0+a\tau) - B_H(t_0);\ \tau \in \mathbb{R}\} \stackrel{d}{=} a^H \{B_H(\tau);\ \tau \in \mathbb{R}\}, \tag{3.2.4}$$

where "$\stackrel{d}{=}$" denotes equality of finite-dimensional distributions. This property implies that $\mathbf{B}_H$ is statistically the same on all time scales. This also implies that the sample paths of $\mathbf{B}_H$ are *fractals*, as defined in [35].

From (3.2.4) it follows that, for any $\delta > 0$, the process $\mathbf{B}_{H,\delta} \triangleq \frac{1}{\delta}\{B_H(t+\delta) - B_H(t);\ t \in \mathbb{R}\}$ is a zero-mean, stationary Gaussian process with covariance function

$$\begin{aligned} K_{\mathbf{B}_{H,\delta}}(\tau) &= \frac{V_H \delta^{2H-2}}{2}\left[\left(\frac{|\tau|}{\delta}+1\right)^{2H} - 2\left(\frac{|\tau|}{\delta}\right)^{2H} + \left|\frac{|\tau|}{\delta}-1\right|^{2H}\right], & \tau \in \mathbb{R}, \\ &\cong V_H H(2H-1)|\tau|^{2H-2}, & |\tau| \gg \delta, \end{aligned} \tag{3.2.5}$$

and power spectral density

$$\begin{aligned} S_{\mathbf{B}_{H,\delta}}(\omega) &= \int_{-\infty}^{\infty} R_{\mathbf{B}_{H,\delta}}(\tau) e^{-i\omega\tau} d\tau, & 0 \neq \omega \in \mathbb{R} \\ &\cong |\omega|^{1-2H}, & 0 < |\omega\delta| \ll 1. \end{aligned} \tag{3.2.6}$$

Equation (3.2.5) implies that the process $\mathbf{B}_{H,\delta}$ is mixing and ergodic but not strongly mixing (in the sense of Rosenblatt [55]). Hence, the increments of FBM provide a simple model for processes with strong long-term dependence. Moreover, (3.2.6) implies that the increments of FBM provide a good model for certain processes with $1/f$-type spectral behavior. Although it is not immediately obvious, (3.2.1) represents a fairly natural way to define a process with this spectral behavior. We present in the following paragraph a heuristic development of the definition of FBM.

In order for all increments to be stationary and exhibit $1/f$-type spectral behavior, $\mathbf{B}_H$ should have a zero-mean, stationary Gaussian "derivative" $\mathbf{W}_H$ with covariance function

$$K_{\mathbf{W}_H}(\tau) = V_H H(2H-1)\,|\tau|^{2H-2}, \qquad \tau \in \mathbb{R},$$

and power spectral density

$$S_{\mathbf{W}_H}(\omega) = |\omega|^{1-2H}, \qquad 0 \neq \omega \in \mathbb{R}.$$

That is, we would like to define $\mathbf{B}_H$ as

$$B_H(t) \triangleq \int_0^t W_H(\tau)\,d\tau, \qquad t \in \mathbb{R}, \tag{3.2.7}$$

where the process $\mathbf{W}_H$ has the properties described above. Following Barnes and Allan [5], we define $\mathbf{W}_H$ to be the $(H-½)^{th}$-order fractional integral of white noise, but since white noise is defined on all of $\mathbb{R}$, we need to define our fractional integrals accordingly. That is, we let

$$W_H(t) \triangleq \frac{1}{\Gamma(H-½)} \int_{-\infty}^{t} |t-\tau|^{H-3/2} W(\tau)\,d\tau, \qquad \forall\, t \in \mathbb{R},$$

where $\mathbf{W}$ is a standard white noise process. Substituting this expression into (3.2.7) yields

$$B_H(t) = \frac{1}{\Gamma(H-½)} \int_0^t \int_{-\infty}^{\tau} |\tau-s|^{H-3/2} W(s)\,ds\,d\tau,$$

which, upon changing the order of integration, becomes

$$B_H(t) = \frac{1}{\Gamma(H-½)} \left[\int_{-\infty}^{0} \left[\int_{0}^{t} |\tau-s|^{H-3/2} d\tau \right] W(s)ds + \int_{0}^{t} \left[\int_{s}^{t} |\tau-s|^{H-3/2} d\tau \right] W(s)ds \right]$$

$$= \frac{1}{\Gamma(H+½)} \left[\int_{-\infty}^{0} (|t-s|^{H-½} - |s|^{H-½}) W(s)ds + \int_{0}^{t} |t-s|^{H-½} W(s)ds \right].$$

Rewriting the white noise integrals in terms of a standard Brownian motion $\mathbf{B}$, we get (3.2.1).

The process $\mathbf{W}_H$ described above is loosely referred to as *fractional Gaussian noise,* or FGN. Clearly, no such process actually exists, but the concept of FGN and its relationship to FBM can be made rigorous by defining $\mathbf{W}_H$ as a generalized Gaussian process; that is, as a linear operator acting on a certain subset of $L^2(\mathbb{R})$. Then, if we let $I_{[0,t]}$ be the indicator function of the interval $[0,t]$, we get

$$B_H(t) = \mathbf{W}_H(I_{[0,t]}), \qquad \forall\, t \in \mathbb{R}. \tag{3.2.8}$$

(Note: for $t < 0$, we let $I_{[0,t]} \triangleq -I_{[t,0]}$.) This is, of course, a generalization of the well-known relationship between Brownian motion and white noise, and one can define FGN in much the same way that one can define white noise, as follows.

Let $\mathbf{B}$ again be a standard Brownian motion, and let $\mathbf{B}_H$ be an FBM derived from $\mathbf{B}$ via (3.2.1). It is well known that there is an orthogonal increments process $\mathbf{Z} \triangleq \{Z(\omega);\ \omega \in \mathbb{R}\}$ with mean zero and variance $1/(2\pi)$ such that, for any function $f \in L^2(\mathbb{R})$ with Fourier transform $\hat{f}$, we have

$$\int_{-\infty}^{\infty} f(t)dB(t) = \int_{-\infty}^{\infty} \hat{f}(-\omega)dZ(\omega), \tag{3.2.9}$$

and $L^2(\mathbf{B}) = L^2(\mathbf{Z})$. One can define standard white noise as an operator $\mathbf{W}$: $L^2(\mathbb{R}) \rightarrow L^2(\mathbf{B})$

by

$$\mathbf{W}(f) \triangleq \int_{-\infty}^{\infty} f(t)dB(t) = \int_{-\infty}^{\infty} \hat{f}(-\omega)dZ(\omega), \qquad f \in L^2(\mathbb{R}).$$

Now, let us define a new process $\mathbf{Z}_{\mathbf{W}_H} \triangleq \{Z_{\mathbf{W}_H}(\omega);\ \omega \in \mathbb{R}\}$ by

$$Z_{\mathbf{W}_H}(\omega) \triangleq \int_0^{\omega} |\lambda|^{\frac{1}{2}-H} e^{-i\,\mathrm{sgn}(\lambda)(H-\frac{1}{2})\frac{\pi}{2}} dZ(\lambda), \qquad \omega \in \mathbb{R}. \tag{3.2.10}$$

It follows (see, for example, [2], Chapter 2) that $\mathbf{Z}_{\mathbf{W}_H}$ is a zero-mean, Gaussian process with orthogonal increments and associated measure μ given by

$$\mu(A) \triangleq \frac{1}{2\pi}\int_A |\omega|^{1-2H} d\omega, \qquad A \text{ Borel.} \tag{3.2.11}$$

Let $\Lambda_H \subset L^2(\mathbb{R})$ be defined as

$$\Lambda_H \triangleq \{g \in L^2(\mathbb{R}) : \hat{g} \in L^2(\mu)\}. \tag{3.2.12}$$

That is, Λ_H is the subset of functions in $L^2(\mathbb{R})$ whose Fourier transforms are square integrable with respect to the measure μ. We define the *fractional Gaussian noise operator* $\mathbf{W}_H : \Lambda_H \to L^2(\mathbf{Z}_{\mathbf{W}_H})$ by

$$\mathbf{W}_H(g) \triangleq \int_{-\infty}^{\infty} \hat{g}(-\omega)dZ_{\mathbf{W}_H}(\omega), \qquad g \in \Lambda_H, \tag{3.2.13}$$

and we will use the notation

$$\int_{-\infty}^{\infty} g(t)W_H(t)dt \triangleq \mathbf{W}_H(g).$$

It follows immediately from (3.2.13) that, for all $f, g \in \Lambda_H$,

$$E\left\{\int_{-\infty}^{\infty} f(t)W_H(t)dt\,\overline{\int_{-\infty}^{\infty} g(s)W_H(s)ds}\right\} = E\left\{\int_{-\infty}^{\infty} \hat{f}(-\omega)dZ_{\mathbf{W}_H}(\omega)\,\overline{\int_{-\infty}^{\infty} \hat{g}(-\omega)dZ_{\mathbf{W}_H}(\omega)}\right\}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(-\omega)\overline{\hat{g}(-\omega)}\,|\omega|^{1-2H}\,d\omega,$$

and

$$E\left\{\int_{-\infty}^{\infty} f(t)W_H(t)dt\right\} = E\left\{\int_{-\infty}^{\infty} g(s)W_H(s)ds\right\} = 0.$$

Also, it can be shown (see Appendix A, Lemmas A.1 and A.2) that $L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \subseteq \Lambda_H$ and, for any $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(-\omega)\overline{\hat{g}(-\omega)}\,|\omega|^{1-2H}\,d\omega = V_H H(2H-1)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\overline{g(s)}\,|t-s|^{2H-2}dsdt.$$

Hence, for any $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, we have

$$E\left\{\int_{-\infty}^{\infty} f(t)W_H(t)dt\,\overline{\int_{-\infty}^{\infty} g(s)W_H(s)ds}\right\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(-\omega)\overline{\hat{g}(-\omega)}\,|\omega|^{1-2H}\,d\omega \tag{3.2.14}$$

$$= V_H H(2H-1)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\overline{g(s)}\,|t-s|^{2H-2}dsdt,$$

so the operator $\mathbf{W}_H$ behaves like a zero-mean, stationary Gaussian process with covariance function $K_{\mathbf{W}_H}$ and power spectral density $S_{\mathbf{W}_H}$, as previously defined.

It remains to be shown that $\mathbf{B}_H$ results from $\mathbf{W}_H$ via (3.2.8); that is,

$$B_H(t) = \int_0^t W_H(\tau)d\tau = \mathbf{W}_H(I_{[0,t]}), \qquad \forall\, t \in \mathbb{R}.$$

To this end, we recall that

$$\begin{aligned}
\int_0^t W_H(\tau)d\tau &= \int_{-\infty}^{\infty} \hat{I}_{[0,t]}(-\omega)dZ_{\mathbf{W}_H}(\omega) \\
&= \int_{-\infty}^{\infty} \hat{I}_{[0,t]}(-\omega)\,|\omega|^{\frac{1}{2}-H} e^{-i\,\mathrm{sgn}(\omega)(H-\frac{1}{2})\frac{\pi}{2}} dZ(\omega) \\
&= \int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega}\,|\omega|^{\frac{1}{2}-H} e^{-i\,\mathrm{sgn}(\omega)(H-\frac{1}{2})\frac{\pi}{2}} dZ(\omega) \\
&= \int_{-\infty}^{\infty} \hat{f}_t(-\omega)dZ(\omega),
\end{aligned}$$

where

$$\hat{f}_t(\omega) \triangleq \frac{1-e^{-i\omega t}}{i\omega}\,|\omega|^{\frac{1}{2}-H} e^{i\,\mathrm{sgn}(\omega)(H-\frac{1}{2})\frac{\pi}{2}}. \tag{3.2.15}$$

Since $\hat{I}_{[0,t]} \in L^2(\mu)$, we have $\hat{f}_t \in L^2(\mathbb{R})$, and it is straightforward to verify that the inverse Fourier transform of $\hat{f}_t$ is given by

$$f_t(\tau) = \frac{1}{\Gamma(H+\frac{1}{2})}\left[I_{(-\infty,0)}(\tau)(|t-\tau|^{H-\frac{1}{2}} - |\tau|^{H-\frac{1}{2}}) + I_{[0,t)}(\tau)\,|t-\tau|^{H-\frac{1}{2}}\right], \qquad \tau \in \mathbb{R}. \tag{3.2.16}$$

It follows that

$$\int_0^t W_H(\tau)d\tau = \int_{-\infty}^{\infty} \hat{f}_t(-\omega)dZ(\omega)$$

$$= \int_{-\infty}^{\infty} f_t(\tau)dB(\tau)$$

$$= \frac{1}{\Gamma(H+\frac{1}{2})}\left[\int_{-\infty}^{0}(|t-\tau|^{H-\frac{1}{2}} - |\tau|^{H-\frac{1}{2}})dB(\tau) + \int_0^t |t-\tau|^{H-\frac{1}{2}}dB(\tau)\right]$$

$$= B_H(t),$$

as desired.

We close this section with a brief discussion concerning stochastic integration with respect to FBM.

Let $\mathbf{B}_H$ be an FBM and let $\{\pi_n\}_{n=1}^{\infty}$ be a sequence of partitions of a compact interval $[a,b]$ with mesh size going to zero; that is,

$$\pi_n = \{t_0, \cdots, t_n : a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b\},$$

where

$$\lim_{n\to\infty} |\pi_n| = \lim_{n\to\infty} \max_{1\le i\le n} (t_i - t_{i-1}) = 0.$$

If f is a bounded function defined on $[a,b]$, we can define the integral of f with respect to $\mathbf{B}_H$, in the usual L^2-fashion, by

$$\int_a^b f(\tau)dB_H(\tau) \triangleq \lim_{n\to\infty} \sum_{i=1}^{n} f(\xi_i)[B_H(t_i) - B_H(t_{i-1})], \qquad \xi_i \in [t_{i-1}, t_i].$$

Of course, the integral is only well-defined if the right-hand limit exists and is the same for all sequences of partitions $\{\pi_n\}_{n=1}^{\infty}$ with $|\pi_n| \to 0$. We can also define "improper" integrals as limits of "proper" ones; for example, if f is bounded and integrable with respect to $\mathbf{B}_H$ on all compact intervals $[-T,T]$, then

$$\int_{-\infty}^{\infty} f(\tau)dB_H(\tau) \triangleq \lim_{T\to\infty} \int_{-T}^{T} f(\tau)dB_H(\tau),$$

whenever the right-hand limit exists.

Now, let $\mathbf{W}_H$ be the FGN corresponding to $\mathbf{B}_H$. We wish to establish some relationships between the operator $\mathbf{W}_H$ and integrals with respect to $\mathbf{B}_H$.

Lemma 3.2.1: If $\{f_n\}_{n=1}^{\infty} \subseteq L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is a sequence of functions that converges in $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$ to a function f, then

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(\tau)W_H(\tau)d\tau = \int_{-\infty}^{\infty} f(\tau)W_H(\tau)d\tau. \tag{3.2.17}$$

Proof: Let μ be the measure defined by (3.2.11). Recall that

$$\mathrm{E}\left\{\left|\int_{-\infty}^{\infty} f_n(\tau)W_H(\tau)d\tau - \int_{-\infty}^{\infty} f(\tau)W_H(\tau)d\tau\right|^2\right\} = \mathrm{E}\left\{\left|\int_{-\infty}^{\infty} [f_n(\tau) - f(\tau)]W_H(\tau)d\tau\right|^2\right\}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}_n(\omega) - \hat{f}(\omega)|^2 |\omega|^{1-2H} d\omega.$$

Hence, to prove (3.2.17), it is sufficient to show that $\hat{f}_n \to \hat{f}$ in $L^2(\mu)$. To this end, notice that, since $f_n \to f$ in $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$, $\hat{f}_n \to \hat{f}$ uniformly and in $L^2(\mathbb{R})$. So, for any $\varepsilon > 0$,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}_n(\omega) - \hat{f}(\omega)|^2 |\omega|^{1-2H} d\omega \le \frac{1}{2\pi}\int_{-1}^{1} |\hat{f}_n(\omega) - \hat{f}(\omega)|^2 |\omega|^{1-2H} d\omega$$

$$+ \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}_n(\omega) - \hat{f}(\omega)|^2 d\omega$$

$$\le \frac{1}{2\pi}\int_{-1}^{1} \varepsilon |\omega|^{1-2H} d\omega + \frac{1}{2\pi}\varepsilon \qquad \text{(for sufficiently large n)}$$

$$= \frac{2-H}{\pi(2-2H)}\varepsilon.$$

Therefore, $\hat{f}_n \to \hat{f}$ in $L^2(\mu)$ and (3.2.17) follows. ■

Lemma 3.2.2: If f is a bounded function, continuous almost everywhere on a compact interval $[a,b]$, then

$$\int_a^b f(\tau)dB_H(\tau) = \int_a^b f(\tau)W_H(\tau)d\tau.$$

Proof: Let $\{\pi_n\}_{n=1}^{\infty}$ be a sequence of partitions of $[a,b]$ with $|\pi_n| \to 0$. Then

$$\int_a^b f(\tau)dB_H(\tau) = \lim_{n\to\infty} \sum_{i=1}^{n} f(\xi_i)[B_H(t_i) - B_H(t_{i-1})]$$

$$= \lim_{n\to\infty} \sum_{i=1}^{n} f(\xi_i)\int_{-\infty}^{\infty} I_{[t_{i-1},t_i]}(\tau)W_H(\tau)d\tau$$

$$= \lim_{n\to\infty} \int_{-\infty}^{\infty} \left[\sum_{i=1}^{n} f(\xi_i)I_{[t_{i-1},t_i]}(\tau)\right]W_H(\tau)d\tau$$

$$= \lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(\tau)W_H(\tau)d\tau,$$

where

$$f_n(\tau) \triangleq \sum_{i=1}^{n} f(\xi_i)I_{[t_{i-1},t_i]}(\tau).$$

Now, since f is continuous almost everywhere on [a,b], $f_n \to fI_{[a,b]}$ almost everywhere on

$\mathbb{R}$, and, since f is bounded, the dominated convergence theorem implies that $f_n \to f I_{[a,b]}$ in $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$. Hence,

$$\begin{aligned}\int_a^b f(\tau)dB_H(\tau) &= \lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(\tau)W_H(\tau)d\tau \\ &= \int_{-\infty}^{\infty} I_{[a,b]}(\tau)f(\tau)W_H(\tau)d\tau \qquad \text{(by Lemma 3.2.1)} \\ &= \int_a^b f(\tau)W_H(\tau)d\tau. \ \blacksquare\end{aligned}$$

Lemma 3.2.3: If $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is bounded and almost everywhere continuous on $\mathbb{R}$, then

$$\int_{-\infty}^{\infty} f(\tau)dB_H(\tau) = \int_{-\infty}^{\infty} f(\tau)W_H(\tau)d\tau.$$

Proof: By definition,

$$\begin{aligned}\int_{-\infty}^{\infty} f(\tau)dB_H(\tau) &= \lim_{N\to\infty} \int_{-N}^{N} f(\tau)dB_H(\tau) \\ &= \lim_{N\to\infty} \int_{-N}^{N} f(\tau)W_H(\tau)d\tau \qquad \text{(by Lemma 3.2.2)} \\ &= \lim_{N\to\infty} \int_{-\infty}^{\infty} f_N(\tau)W_H(\tau)d\tau,\end{aligned}$$

where

$$f_N(\tau) \triangleq f(\tau)I_{[-N,N]}(\tau).$$

Clearly, $f_N \to f$ in $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$, so

$$\int_{-\infty}^{\infty} f(\tau)dB_H(\tau) = \lim_{N\to\infty} \int_{-\infty}^{\infty} f_N(\tau)W_H(\tau)d\tau$$

$$= \int_{-\infty}^{\infty} f(\tau)W_H(\tau)d\tau \qquad \text{(by Lemma 3.2.1).} \ \blacksquare$$

These lemmas are, of course, not exhaustive, but they illustrate that, for sufficiently well-behaved functions f, the random variable $\mathbf{W}_H(f)$ is equivalent to the integral of f with respect to $\mathbf{B}_H$.

3.3. RKHS Results for Fractional Brownian Motion

In this section, we characterize the RKHSs for the class of FBMs and present some related results. In particular, we develop several results concerning the restriction of an FBM to a compact index set $[0,T]$, which will be directly applicable to the problem of detecting signals in FGN. We begin by considering the unrestricted FBM process, in which the index set is $\mathbb{R}$. For this case, we have the following result.

Theorem 3.3.1: Let $\mathbf{B}_H$ be an FBM and let μ be the measure defined by (3.2.11). Then $\mathcal{H}(\mathbf{B}_H)$ consists of functions of the form

$$g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega}\overline{\tilde{g}(\omega)}|\omega|^{1-2H}d\omega, \qquad t \in \mathbb{R}, \tag{3.3.1}$$

where $\tilde{g} \in L^2(\mu)$. Further, for all $g, h \in \mathcal{H}(\mathbf{B}_H)$,

$$\langle g, h\rangle_{\mathcal{H}(\mathbf{B}_H)} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{h}(\omega)\overline{\tilde{g}(\omega)}|\omega|^{1-2H}d\omega. \tag{3.3.2}$$

Proof: Let $\mathbf{W}_H$ be the FGN corresponding to $\mathbf{B}_H$ and let $\mathbf{Z}_{\mathbf{W}_H}$ be the orthogonal increments process related to $\mathbf{W}_H$ by (3.2.13). Then

$$
\begin{aligned}
K_{\mathbf{B}_H}(s,t) &= \mathrm{E}\{B_H(s)\overline{B_H(t)}\} \\
&= \mathrm{E}\left\{\int_0^s W_H(\tau)d\tau \overline{\int_0^t W_H(\tau)d\tau}\right\} \\
&= \mathrm{E}\left\{\int_{-\infty}^{\infty} \hat{I}_{[0,s]}(-\omega)dZ_{\mathbf{W}_H}(\omega) \overline{\int_{-\infty}^{\infty} \hat{I}_{[0,t]}(-\omega)dZ_{\mathbf{W}_H}(\omega)}\right\} \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{I}_{[0,s]}(-\omega)\overline{\hat{I}_{[0,t]}(-\omega)}\,|\omega|^{1-2H}\,d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega s}-1}{i\omega}\overline{\frac{e^{i\omega t}-1}{i\omega}}\,|\omega|^{1-2H}\,d\omega.
\end{aligned}
$$

It can be shown (see Appendix A, Lemma A.3) that the set of functions $\{\frac{e^{i\omega t}-1}{i\omega};\ t \in \mathbb{R}\}$ spans $L^2(\mu)$, so (3.3.1) and (3.3.2) follow by applying Theorem 2.1.1. ■

Theorem 3.3.1 gives a frequency-domain characterization of $\boldsymbol{H}(\mathbf{B}_H)$. The following corollary gives the corresponding time-domain characterization.

Corollary 3.3.2: A function $g \in \boldsymbol{H}(\mathbf{B}_H)$ if and only if there exists $g^* \in L^2(\mathbb{R})$ such that

$$
g(t) = \frac{1}{\Gamma(H-\frac{1}{2})}\int_0^t\int_{-\infty}^s (s-\tau)^{H-\frac{3}{2}}\overline{g^*(\tau)}\,d\tau ds, \qquad \forall\, t \in \mathbb{R}. \tag{3.3.3}
$$

Further, for all $g,h \in \boldsymbol{H}(\mathbf{B}_H)$,

$$
\langle g, h\rangle_{\boldsymbol{H}(\mathbf{B}_H)} = \int_{-\infty}^{\infty} h^*(s)\overline{g^*(s)}ds. \tag{3.3.4}
$$

Proof: By Theorem 3.3.1, $g \in \boldsymbol{H}(\mathrm{B}_H)$ if and only if there exists $\tilde{g} \in L^2(\mu)$ such that, for all $t \in \mathbb{R}$,

$$g(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega}\overline{\tilde{g}(\omega)}\,|\omega|^{1-2H}\,d\omega$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\frac{1-e^{-i\omega t}}{i\omega}|\omega|^{\frac{1}{2}-H}e^{i\,\mathrm{sgn}(\omega)(H-\frac{1}{2})\frac{\pi}{2}}\right]\left[\overline{\tilde{g}(-\omega)\,|\omega|^{\frac{1}{2}-H}e^{i\,\mathrm{sgn}(\omega)(H-\frac{1}{2})\frac{\pi}{2}}}\right]d\omega$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}_t(\omega)\overline{\hat{g}^*(\omega)}\,d\omega,$$

where

$$\hat{f}_t(\omega) \triangleq \frac{1-e^{-i\omega t}}{i\omega}|\omega|^{\frac{1}{2}-H}e^{i\,\mathrm{sgn}(\omega)(H-\frac{1}{2})\frac{\pi}{2}} \qquad \text{(cf. (3.2.15))},$$

and

$$\hat{g}^*(\omega) \triangleq \tilde{g}(-\omega)\,|\omega|^{\frac{1}{2}-H}e^{i\,\mathrm{sgn}(\omega)(H-\frac{1}{2})\frac{\pi}{2}}.$$

Clearly, $\hat{f}_t, \hat{g}^* \in L^2(\mathbb{R})$, so let g^* denote the inverse Fourier transform of $\hat{g}^*$ and recall from (3.2.16) that the inverse Fourier transform of $\hat{f}_t$ is given by

$$f_t(\tau) = \frac{1}{\Gamma(H+\frac{1}{2})}\left[I_{(-\infty,0)}(\tau)(|t-\tau|^{H-\frac{1}{2}} - |\tau|^{H-\frac{1}{2}}) + I_{[0,t)}(\tau)|t-\tau|^{H-\frac{1}{2}}\right], \qquad \tau \in \mathbb{R}.$$

Hence,

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}_t(\omega)\overline{\hat{g}^*(\omega)}d\omega$$

$$= \int_{-\infty}^{\infty} f_t(\tau)\overline{g^*(\tau)}d\tau$$

$$= \frac{1}{\Gamma(H+½)}\left[\int_{-\infty}^{0} (|t-\tau|^{H-½} - |\tau|^{H-½})\overline{g^*(\tau)}d\tau + \int_{0}^{t} |t-\tau|^{H-½}\overline{g^*(\tau)}d\tau\right]$$

$$= \frac{1}{\Gamma(H-½)} \int_{0}^{t}\int_{-\infty}^{s} (s-\tau)^{H-3/2}\overline{g^*(\tau)}d\tau ds,$$

which proves (3.3.3).

To prove (3.3.4), we proceed in a similar fashion. Let $g, h \in \boldsymbol{H}(\mathrm{B}_H)$ and let $\tilde{g}, \tilde{h} \in L^2(\mu)$ be given by Theorem 3.3.1. Then (3.3.2) yields

$$\langle g, h\rangle_{\boldsymbol{H}(\mathrm{B}_H)} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{h}(\omega)\overline{\tilde{g}(\omega)}|\omega|^{1-2H}d\omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\tilde{h}(-\omega)|\omega|^{½-H} e^{i\,\mathrm{sgn}(\omega)(H-½)\frac{\pi}{2}}\right]\left[\overline{\tilde{g}(-\omega)|\omega|^{½-H} e^{i\,\mathrm{sgn}(\omega)(H-½)\frac{\pi}{2}}}\right]d\omega$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{h}^*(\omega)\overline{\hat{g}^*(\omega)}d\omega \qquad (\hat{h}^*, \hat{g}^* \text{ defined as before})$$

$$= \int_{-\infty}^{\infty} h^*(\tau)\overline{g^*(\tau)}d\tau,$$

which proves (3.3.4). ■

Remark: Notice that (3.3.3) implies that any function $g \in \boldsymbol{H}(\mathrm{B}_H)$ is differentiable almost everywhere in $\boldsymbol{R}$ and that the derivative g' is the $(H-½)^{th}$–order fractional integral of $\overline{g^*}$. That is, for almost all $t \in \boldsymbol{R}$,

$$g'(t) = \frac{1}{\Gamma(H-\frac{1}{2})} \int_{-\infty}^{t} (t-\tau)^{H-\frac{3}{2}} \overline{g^*(\tau)} d\tau. \tag{3.3.5}$$

Functions of the form (3.3.5) are variants of the so-called *Riesz potentials*, which have been widely studied (see, for example, [59]). This representation can be exploited to give various conditions necessary for $g \in \boldsymbol{H}(\mathbf{B}_H)$. For example, it follows from (3.3.5) and a general form of Young's inequality (see [15], page 232) that $g' \in L^{\frac{1}{1-H}}(\mathbb{R})$. Similarly, from (3.3.1) we can get a simple sufficient condition for $g \in \boldsymbol{H}(\mathbf{B}_H)$. Letting $\hat{g}'$ be the Fourier transform of g', it follows from (3.3.1) that $g \in \boldsymbol{H}(\mathbf{B}_H)$ if $\hat{g}' \in L^1(\mathbb{R})$ and

$$\int_{-\infty}^{\infty} |\hat{g}'(\omega)|^2 |\omega|^{2H-1} d\omega < \infty.$$

The following theorem gives an infinite interval *whitening filter* for FBM.

Theorem 3.3.3: Let $\mathbf{B}$ be a standard Brownian motion and let $\mathbf{B}_H$ be an FBM derived from $\mathbf{B}$ via (3.2.1). Then

$$B(t) = \frac{1}{\Gamma(\frac{3}{2}-H)} \left[\int_{-\infty}^{0} (|t-\tau|^{\frac{1}{2}-H} - |\tau|^{\frac{1}{2}-H}) dB_H(\tau) + \int_{0}^{t} |t-\tau|^{\frac{1}{2}-H} dB_H(\tau) \right]. \tag{3.3.6}$$

Proof: Again, let $\mathbf{W}_H$ be the FGN corresponding to $\mathbf{B}_H$, $\mathbf{Z}_{\mathbf{W}_H}$ the associated orthogonal increments process, and μ the measure given by (3.2.11). Also, let $\mathbf{Z}$ be the orthogonal increments process related to $\mathbf{B}$ by (3.2.9). Then, for all $t \in \mathbb{R}$,

$$
\begin{aligned}
B(t) &= \int_0^t dB(t) \\
&= \int_{-\infty}^{\infty} \hat{I}_{[0,t]}(-\omega)\,dZ(\omega) \\
&= \int_{-\infty}^{\infty} \hat{I}_{[0,t]}(-\omega)\,|\omega|^{H-\frac{1}{2}} e^{i\,\mathrm{sgn}(\omega)(H-\frac{1}{2})\frac{\pi}{2}}\,dZ_{W_H}(\omega) \qquad \text{(by inverting (3.2.10))} \\
&= \int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega}\,|\omega|^{H-\frac{1}{2}} e^{i\,\mathrm{sgn}(\omega)(H-\frac{1}{2})\frac{\pi}{2}}\,dZ_{W_H}(\omega) \\
&= \int_{-\infty}^{\infty} \hat{f}_t(-\omega)\,dZ_{W_H}(\omega),
\end{aligned}
$$

where

$$
\hat{f}_t(\omega) \triangleq \frac{1-e^{-i\omega t}}{i\omega}\,|\omega|^{H-\frac{1}{2}} e^{-i\,\mathrm{sgn}(\omega)(H-\frac{1}{2})\frac{\pi}{2}}.
$$

Clearly, $\hat{f}_t \in L^2(\mu)$, and it is straightforward to show that $\hat{f}_t \in L^2(\mathbb{R})$ with inverse Fourier transform $f_t \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ given by

$$
f_t(\tau) = \frac{1}{\Gamma(\frac{3}{2}-H)} \left[I_{(-\infty,0)}(\tau)\left(|t-\tau|^{\frac{1}{2}-H} - |\tau|^{\frac{1}{2}-H}\right) + I_{[0,t)}(\tau)\,|t-\tau|^{\frac{1}{2}-H} \right], \qquad \tau \in \mathbb{R}.
$$

If we let

$$
f_N(\tau) \triangleq \left[1 - I_{[-\frac{1}{N},\frac{1}{N}]}(\tau)\right]\left[1 - I_{[t-\frac{1}{N},t+\frac{1}{N}]}(\tau)\right] I_{[-N,N]}(\tau) f_t(\tau), \qquad \tau \in \mathbb{R},
$$

then $f_N \to f_t$ in $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$, and

$$
\begin{aligned}
B(t) &= \int_{-\infty}^{\infty} \hat{f}_t(-\omega)\, dZ_{W_H}(\omega) \\
&= \int_{-\infty}^{\infty} f_t(\tau) W_H(\tau)\, d\tau \\
&= \lim_{N\to\infty} \int_{-\infty}^{\infty} f_N(\tau) W_H(\tau)\, d\tau \qquad \text{(by lemma 3.2.1)} \\
&= \lim_{N\to\infty} \int_{-\infty}^{\infty} f_N(\tau)\, dB_H(\tau) \qquad \text{(by lemma 3.2.3)} \\
&= \int_{-\infty}^{\infty} f_t(\tau)\, dB_H(\tau) \\
&= \frac{1}{\Gamma(3/2-H)} \left[\int_{-\infty}^{0} (|t-\tau|^{1/2-H} - |\tau|^{1/2-H})\, dB_H(\tau) + \int_0^t |t-\tau|^{1/2-H}\, dB_H(\tau) \right]. \quad \blacksquare
\end{aligned}
$$

We now consider the case of an FBM restricted to a compact index set $[0,T]$. Let $\mathbf{B}_H$ be an FBM and let $\mathbf{B}_{H|T} \triangleq \{B_{H|T}(t);\, t \in [0,T]\}$ represent the restriction of $\mathbf{B}_H$ to $[0,T]$. It is well-known that $\boldsymbol{H}(\mathbf{B}_{H|T})$ must consist entirely of functions in $\boldsymbol{H}(\mathbf{B}_H)$ restricted to the interval $[0,T]$. Unfortunately, this is not a very practical characterization, since it is typically rather difficult to determine from its values on $[0,T]$ whether a particular function can be extended to all of $\boldsymbol{R}$ in such a way that the extension is a member of $\boldsymbol{H}(\mathbf{B}_H)$. Fortunately, there is a more useful description of $\boldsymbol{H}(\mathbf{B}_{H|T})$, as the following theorem shows.

Theorem 3.3.4: $\boldsymbol{H}(\mathbf{B}_{H|T})$ consists of functions of the form

$$
g(t) = \frac{1}{\Gamma(H-1/2)} \int_0^t \left[\int_s^t \tau^{H-1/2} (\tau-s)^{H-3/2}\, d\tau \right] \overline{\tilde{g}(s)}\, s^{1/2-H}\, ds, \qquad t \in [0,T], \tag{3.3.7}
$$

where $\tilde{g} \in L^2([0,T])$. Any such function g has a derivative g' almost everywhere in [0,T], and, for almost all $t \in [0,T]$,

$$\overline{\tilde{g}(t)} = t^{H-\frac{1}{2}} \frac{d}{dt}\left[\frac{1}{\Gamma(\frac{3}{2}-H)}\int_0^t (t-s)^{\frac{1}{2}-H} s^{\frac{1}{2}-H} g'(s)\,ds\right]. \tag{3.3.8}$$

Further, for any $g, h \in \boldsymbol{H}(\mathbf{B}_{H\,|T})$,

$$\langle g, h\rangle_{\boldsymbol{H}(\mathbf{B}_{H\,|T})} = \int_0^T \tilde{h}(\tau)\overline{\tilde{g}(\tau)}\,d\tau. \tag{3.3.9}$$

Proof: Since $\mathbf{B}_{H\,|T}$ is the restriction of $\mathbf{B}_H$ to $[0,T]$, we have, for all $s,t \in [0,T]$,

$$\begin{aligned}
K_{\mathbf{B}_{H|T}}(s,t) &= K_{\mathbf{B}_H}(s,t) \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{I}_{[0,s]}(-\omega)\overline{\hat{I}_{[0,t]}(-\omega)}\,|\omega|^{1-2H}\,d\omega \\
&= V_H H(2H-1)\int_0^T\!\!\int_0^T I_{[0,s]}(\sigma) I_{[0,t]}(\tau)\,|\tau-\sigma|^{2H-2}\,d\tau\,d\sigma \qquad \text{(by (3.2.14))} \\
&= \frac{-\Gamma(2-2H)\cos\pi H}{\pi}\int_0^s\!\!\int_0^t |\tau-\sigma|^{2H-2}\,d\tau\,d\sigma \qquad \text{(by (3.2.3))}.
\end{aligned}$$

The function $|\tau-\sigma|^{2H-2}$ can be decomposed as (see [75], Equation (15))

$$|\tau-\sigma|^{2H-2} = \frac{\Gamma(\frac{3}{2}-H)}{\Gamma(H-\frac{1}{2})\Gamma(2-2H)}(\sigma\tau)^{H-\frac{1}{2}} \int_0^{\min(\sigma,\tau)} u^{1-2H}(\sigma-u)^{H-\frac{3}{2}}(\tau-u)^{H-\frac{3}{2}}\,du.$$

Hence,

$$K_{B_{H|T}}(s,t) = \frac{-\Gamma(2-2H)\cos\pi H}{\pi} \int_0^s\int_0^t |\tau-\sigma|^{2H-2} d\tau d\sigma$$

$$= \frac{-\Gamma(3/2-H)\cos\pi H}{\Gamma(H-1/2)\pi} \int_0^s\int_0^t (\sigma\tau)^{H-1/2} \left[\int_0^{\min(\sigma,\tau)} u^{1-2H} (\sigma-u)^{H-3/2} (\tau-u)^{H-3/2} du \right] d\tau d\sigma$$

$$= \left[\frac{1}{\Gamma(H-1/2)}\right]^2 \int_0^{\min(s,t)} u^{1-2H} \left[\int_u^s \sigma^{H-1/2}(\sigma-u)^{H-3/2} d\sigma\right] \left[\int_u^t \tau^{H-1/2}(\tau-u)^{H-3/2} d\tau\right] du$$

$$= \int_0^T f_s(u) f_t(u) du,$$

where

$$f_t(u) \triangleq \frac{1}{\Gamma(H-1/2)} I_{[0,t]}(u) u^{1/2-H} \int_u^t \tau^{H-1/2}(\tau-u)^{H-3/2} d\tau. \qquad (3.3.10)$$

It can be shown (see Appendix A, Lemma A.4) that the functions $\{f_t;\, t \in [0,T]\}$ span $L^2([0,T])$, so (3.3.7) and (3.3.9) follow from Theorem 2.1.1. To establish (3.3.8), we choose an arbitrary function $g \in \boldsymbol{H}(B_{H|T})$ and use (3.3.7) to write

$$g(t) = \frac{1}{\Gamma(H-1/2)} \int_0^t \left[\int_s^t \tau^{H-1/2}(\tau-s)^{H-3/2} d\tau\right] \overline{g(s)} s^{1/2-H} ds$$

$$= \frac{1}{\Gamma(H-1/2)} \int_0^t \tau^{H-1/2} \left[\int_0^\tau (\tau-s)^{H-3/2} s^{1/2-H}\, \overline{g(s)} ds\right] d\tau.$$

It follows that, for almost all $t \in [0,T]$,

$$g'(t) = \frac{1}{\Gamma(H-1/2)} t^{H-1/2} \int_0^t (t-s)^{H-3/2} s^{1/2-H}\, \overline{g(s)} ds.$$

Hence,

$$\frac{1}{\Gamma(3/2-H)}\int_0^t (t-s)^{1/2-H}s^{1/2-H}g'(s)ds = \frac{1}{\Gamma(3/2-H)}\int_0^t (t-s)^{1/2-H}\left[\frac{1}{\Gamma(H-1/2)}\int_0^s (s-u)^{H-3/2}u^{1/2-H}\overline{\overline{g}(u)}du\right]ds$$

$$= \frac{1}{\Gamma(3/2-H)\Gamma(H-1/2)}\int_0^t u^{1/2-H}\overline{\overline{g}(u)}\left[\int_u^t (t-s)^{1/2-H}(s-u)^{H-3/2}ds\right]du$$

$$= \int_0^t u^{1/2-H}\overline{\overline{g}(u)}du.$$

Therefore, for almost all $t \in [0,T]$,

$$t^{H-1/2}\frac{d}{dt}\left[\frac{1}{\Gamma(3/2-H)}\int_0^t (t-s)^{1/2-H}s^{1/2-H}g'(s)ds\right] = t^{H-1/2}\frac{d}{dt}\int_0^t u^{1/2-H}\overline{\overline{g}(u)}du$$

$$= \overline{\overline{g}(t)},$$

which proves (3.3.8). ■

We can also find a canonical representation for $\mathbf{B}_{H|T}$, different from (3.2.1), which leads straightforwardly to a finite-interval whitening filter, as follows.

Theorem 3.3.5: Let $\{f_t;\ t \in [0,T]\}$ be as defined in (3.3.10). There exists a process $\mathbf{B}_T \triangleq \{B_T(t);\ t \in [0,T]\}$, which is a standard Brownian motion on $[0,T]$, such that, for all $t \in [0,T]$,

$$B_{H|T}(t) = \int_0^T f_t(u)dB_T(u)$$
$$= \frac{1}{\Gamma(H-1/2)}\int_0^t u^{1/2-H}\left[\int_u^t \tau^{H-1/2}(\tau-u)^{H-3/2}d\tau\right]dB_T(u), \tag{3.3.11}$$

and

$$B_T(t) = \frac{1}{\Gamma(3/2-H)} \int_0^t \tau^{H-1/2} d_\tau \int_0^\tau (\tau-u)^{1/2-H} u^{1/2-H} \, dB_{H|T}(u), \tag{3.3.12}$$

where the integral in (3.3.12) is to be interpreted as an integral with respect to the orthogonal increments process $\mathbf{Z}_T \triangleq \{Z_T(t);\ t \in [0,T]\}$ given by

$$Z_T(t) = \frac{1}{\Gamma(3/2-H)} \int_0^t (t-\tau)^{1/2-H} \tau^{1/2-H} \, dB_{H|T}(\tau), \qquad t \in [0,T]. \tag{3.3.13}$$

Proof: From the proof of Theorem 3.3.4, we know that

$$K_{\mathbf{B}_{H|T}}(s,t) = \int_0^T f_s(u) f_t(u) du, \qquad \forall\ s,t \in [0,T].$$

It follows from Theorem 2.1.2 that there exists a zero-mean, orthogonal increments process $\mathbf{B}_T \triangleq \{B_T(t);\ t \in [0,T]\}$ with associated measure Lebesgue measure such that

$$B_{H|T}(t) = \int_0^T f_t(u) dB_T(u).$$

Since the functions $\{f_t;\ t \in [0,T]\}$ span $L^2([0,T])$, Theorem 2.1.2 also guarantees that $L^2(\mathbf{B}_{H|T}) = L^2(\mathbf{B}_T)$, which implies that $\mathbf{B}_T$ is Gaussian. Hence, $\mathbf{B}_T$, is a zero-mean, Gaussian, orthogonal increments process on $[0,T]$ with associated measure Lebesgue measure. It follows that $\mathbf{B}_T$ is a standard Brownian motion on $[0,T]$, which proves (3.3.11).

To prove (3.3.12), we define a new orthogonal increments process $\mathbf{Z}_T \triangleq \{Z_T(t);\ t \in [0,T]\}$ by

$$Z_T(t) \triangleq \int_0^t \tau^{1/2-H} dB_T(\tau), \qquad t \in [0,T]. \tag{3.3.14}$$

This definition implies that

$$B_T(t) = \int_0^t \tau^{H-1/2} dZ_T(\tau), \qquad \forall\, t \in [0,T], \tag{3.3.15}$$

so (3.3.12) will be proven if we can establish (3.3.13); that is, if we can show that

$$Z_T(t) = \frac{1}{\Gamma(3/2-H)} \int_0^t (t-\tau)^{1/2-H}\, \tau^{1/2-H}\, dB_{H\,|T}(\tau), \qquad \forall\, t \in [0,T].$$

To this end, let $\{\pi_n\}_{n=1}^{\infty}$ be a sequence of partitions of the interval $[0,T]$ with $|\pi_n| \to 0$. Then, for all $t \in [0,T]$,

$$\frac{1}{\Gamma(3/2-H)} \int_0^t (t-\tau)^{1/2-H}\, \tau^{1/2-H}\, dB_{H\,|T}(\tau) = \lim_{\varepsilon \to 0} \frac{1}{\Gamma(3/2-H)} \int_0^T I_{[\varepsilon,t-\varepsilon]}(\tau)(t-\tau)^{1/2-H}\, \tau^{1/2-H}\, dB_{H\,|T}(\tau)$$

$$= \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{\Gamma(3/2-H)} \sum_{i=1}^{n} I_{[\varepsilon,t-\varepsilon]}(\xi_i)(t-\xi_i)^{1/2-H}\, \xi_i^{1/2-H} [B_{H\,|T}(\tau_i) - B_{H\,|T}(\tau_{i-1})]$$

$$= \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{\Gamma(3/2-H)} \sum_{i=1}^{n} I_{[\varepsilon,t-\varepsilon]}(\xi_i)(t-\xi_i)^{1/2-H}\, \xi_i^{1/2-H} \int_0^T [f_{\tau_i}(u) - f_{\tau_{i-1}}(u)]\, dB_T(u) \qquad \text{(by (3.3.11))}$$

$$= \lim_{\varepsilon \to 0} \lim_{n \to \infty} \frac{1}{\Gamma(3/2-H)} \int_0^T \left\{ \sum_{i=1}^{n} I_{[\varepsilon,t-\varepsilon]}(\xi_i)(t-\xi_i)^{1/2-H}\, \xi_i^{1/2-H} [f_{\tau_i}(u) - f_{\tau_{i-1}}(u)] \right\} dB_T(u)$$

$$= \lim_{\varepsilon \to 0} \frac{1}{\Gamma(3/2-H)} \int_0^T \left\{ \lim_{n \to \infty} \sum_{i=1}^{n} I_{[\varepsilon,t-\varepsilon]}(\xi_i)(t-\xi_i)^{1/2-H}\, \xi_i^{1/2-H} [f_{\tau_i}(u) - f_{\tau_{i-1}}(u)] \right\} dB_T(u)$$

$$= \lim_{\varepsilon \to 0} \frac{1}{\Gamma(3/2-H)} \int_0^T \left[\int_{\varepsilon}^{t-\varepsilon} (t-\tau)^{1/2-H}\, \tau^{1/2-H}\, d_\tau f_\tau(u) \right] dB_T(u)$$

$$= \frac{1}{\Gamma(3/2-H)} \int_0^T \left[\int_0^t (t-\tau)^{1/2-H}\, \tau^{1/2-H}\, d_\tau f_\tau(u) \right] dB_T(u).$$

Now, it is straightforward to show that

$$\frac{1}{\Gamma(3/2-H)} \int_0^t (t-\tau)^{1/2-H}\, \tau^{1/2-H}\, d_\tau f_\tau(u) = I_{[0,t]}(u)\, u^{1/2-H},$$

whence it follows that

$$\frac{1}{\Gamma(3/2-H)}\int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H}\,dB_{H\,|T}(\tau) = \frac{1}{\Gamma(3/2-H)}\int_0^T\left[\int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H}\,d_\tau f_\tau(u)\right]dB_T(u)$$

$$= \int_0^T I_{[0,t]}(u)\,u^{1/2-H}\,dB_T(u)$$

$$= \int_0^t u^{1/2-H}\,dB_T(u)$$

$$= Z_T(t) \qquad \text{(by (3.3.14))}.$$

This establishes (3.3.13) and proves the theorem. ■

The previous two theorems admit the following useful corollary.

Corollary 3.3.6: Given any $g \in \mathbf{H}(\mathbf{B}_{H\,|T})$, the corresponding element in $L^2(\mathbf{B}_{H\,|T})$ is given by

$$\langle \mathbf{B}_{H\,|T}, g\rangle_{\mathbf{H}(\mathbf{B}_{H\,|T})} = \left[\frac{1}{\Gamma(3/2-H)}\right]^2 \int_0^T t^{2H-1}\left[\frac{d}{dt}\int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H}g'(\tau)d\tau\right] d_t\int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H}\,dB_{H\,|T}(\tau). \tag{3.3.16}$$

Proof: Let B_T and Z_T be as in the proof of Theorem 3.3.5. It follows from Theorem 2.1.2 that

$$\langle \mathbf{B}_{H\,|T}, g\rangle_{\mathbf{H}(\mathbf{B}_{H\,|T})} = \int_0^T \tilde{g}(t)\,dB_T(t),$$

where $\tilde{g} \in L^2([0,T])$ is given by Theorem 3.3.4. Hence,

$$\left\langle \mathrm{B}_{H|T}, g \right\rangle_{\mathbf{H}(\mathrm{B}_{H|T})} = \int_0^T \bar{g}(t)\,dB_T(t)$$

$$= \frac{1}{\Gamma(3/2-H)} \int_0^T t^{H-1/2} \left[\frac{d}{dt} \int_0^t (t-\tau)^{1/2-H} \tau^{1/2-H} g'(\tau)\,d\tau \right] dB_T(t) \qquad \text{(by (3.3.8))}$$

$$= \frac{1}{\Gamma(3/2-H)} \int_0^T t^{2H-1} \left[\frac{d}{dt} \int_0^t (t-\tau)^{1/2-H} \tau^{1/2-H} g'(\tau)\,d\tau \right] dZ_T(t) \qquad \text{(by (3.3.15))}$$

$$= \left[\frac{1}{\Gamma(3/2-H)} \right]^2 \int_0^T t^{2H-1} \left[\frac{d}{dt} \int_0^t (t-\tau)^{1/2-H} \tau^{1/2-H} g'(\tau)\,d\tau \right] d_t \int_0^t (t-\tau)^{1/2-H} \tau^{1/2-H}\, dB_{H|T}(\tau) \qquad \text{(by (3.3.13))}.$$

which proves (3.3.16). ■

Remark: The RKHS $\mathbf{H}(\mathrm{B}_{H|T})$ was also studied in an early paper by Molcan and Golosov [38], in which it is claimed that $f \in \mathbf{H}(\mathrm{B}_{H|T})$ if and only if $f' \in L^{\frac{1}{1-H}}([0,T])$. This is certainly a necessary condition (as discussed in the remark following Corollary 3.3.2) but does not seem to be sufficient. Molcan and Golosov also give a canonical representation for $\mathrm{B}_{H|T}$, which is essentially equivalent to (3.3.11) and (3.3.12); however, the result is presented without proof.

3.4. Detection of Deterministic Signals in FGN

In this section, we consider the problem of detecting a deterministic signal in the presence of additive FGN on a compact observation interval $[0,T]$. This corresponds to the following hypothesis testing problem:

$$H_0:\ dY(t) = dB_{H|T}(t), \qquad t \in [0,T]$$

versus (3.4.1)

$$H_1:\ dY(t) = s(t)dt + dB_{H|T}(t), \qquad t \in [0,T],$$

where $\mathrm{Y} \triangleq \{Y(t);\ t \in [0,T]\}$ is the observed process, $\mathrm{B}_{H|T}$ is an FBM restricted to $[0,T]$, and s is a real-valued function defined on $[0,T]$. Problem (3.4.1) can be stated more rigorously as

$$H_0:\ Y(t) = B_{H|T}(t), \qquad t \in [0,T]$$

versus (3.4.2)

$$H_1:\ Y(t) = m(t) + B_{H|T}(t), \qquad t \in [0,T],$$

where

$$m(t) \triangleq \int_0^t s(\tau)d\tau. \tag{3.4.3}$$

We would like to know

(i) under what conditions on the signal m is Problem (3.4.2) nonsingular, and

(ii) in the event that (3.4.2) is nonsingular, what is the formula for the likelihood ratio?

We can make use of the results from the preceding section to answer (i) and (ii), as follows.

Theorem 3.4.1: Problem (3.4.2) is nonsingular if and only if

$$m(t) = \frac{1}{\Gamma(H-\frac{1}{2})}\int_0^t \left[\int_\sigma^t \tau^{H-\frac{1}{2}}(\tau-\sigma)^{H-\frac{3}{2}}d\tau\right]\tilde{m}(\sigma)\sigma^{\frac{1}{2}-H}\,d\sigma, \qquad \forall\ t \in [0,T], \tag{3.4.4}$$

where $\tilde{m} \in L^2([0,T])$, and, for almost all $t \in [0,T]$,

$$\tilde{m}(t) = t^{H-\frac{1}{2}} \frac{d}{dt} \left[\frac{1}{\Gamma(\frac{3}{2}-H)} \int_0^t (t-\tau)^{\frac{1}{2}-H} \tau^{\frac{1}{2}-H} m'(\tau) d\tau \right]. \tag{3.4.5}$$

Proof: Since m is real-valued, Y is Gaussian with covariance function $K_Y = K_{B_{H|T}}$ under both hypotheses, so Problem (3.4.2) reduces to

H_0: Y has mean zero

versus

H_1: Y has mean function m.

Hence, by Theorem 2.1.3, (3.4.2) is nonsingular if and only if $m \in \boldsymbol{H}(Y)$. Since $K_Y = K_{B_{H|T}}$, we have $\boldsymbol{H}(Y) = \boldsymbol{H}(B_{H|T})$, so (3.4.4) and (3.4.5) follow immediately from Theorem 4.4. ■

Theorem 3.4.2: If m satisfies (3.4.4) so that Problem (3.4.2) is nonsingular, then the likelihood ratio is given by

$$L(Y) = \exp\left\{ \langle Y, m \rangle_{\boldsymbol{H}(Y)} - \tfrac{1}{2} \langle m, m \rangle_{\boldsymbol{H}(Y)} \right\}, \tag{3.4.6}$$

where

$$\langle m, m \rangle_{\boldsymbol{H}(Y)} = \left[\frac{1}{\Gamma(\frac{3}{2}-H)} \right]^2 \int_0^T t^{2H-1} \left[\frac{d}{dt} \int_0^t (t-\tau)^{\frac{1}{2}-H} \tau^{\frac{1}{2}-H} m'(\tau) d\tau \right]^2 dt, \tag{3.4.7}$$

and

$$\langle \mathrm{Y}, m\rangle_{\boldsymbol{H}(\mathrm{Y})} = \left[\frac{1}{\Gamma(3/2-H)}\right]^2 \int_0^T t^{2H-1}\left[\frac{d}{dt}\int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H} m'(\tau)d\tau\right] d_t \int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H}\, dY(\tau). \tag{3.4.8}$$

Proof: If m satisfies (3.4.4), then $m \in \boldsymbol{H}(\mathrm{Y}) = \boldsymbol{H}(\mathrm{B}_{H\mid T})$, and (3.4.6) and (3.4.7) follow immediately from Theorems 2.1.3 and 3.3.4, respectively. To prove (3.4.8), we will show that for any function $g \in \boldsymbol{H}(\mathrm{Y})$, $\langle \mathrm{Y}, g\rangle_{\boldsymbol{H}(\mathrm{Y})}$ is given by

$$\langle \mathrm{Y}, g\rangle_{\boldsymbol{H}(\mathrm{Y})} = \phi(g)$$

$$\triangleq \left[\frac{1}{\Gamma(3/2-H)}\right]^2 \int_0^T t^{2H-1}\left[\overline{\frac{d}{dt}\int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H} g'(\tau)d\tau}\right] d_t \int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H}\, dY(\tau). \tag{3.4.9}$$

Equation (3.4.8) then follows by letting $g = m$. To prove (3.4.9), it is sufficient to show that (cf. (2.1.1))

$$\begin{aligned} g(t) &= \mathrm{Cov}\{Y(t), \phi(g)\} \\ &= \mathrm{E}\left\{[Y(t)-\mathrm{E}\{Y(t)\}]\overline{[\phi(g)-\mathrm{E}\{\phi(g)\}]}\right\} \\ &= \mathrm{E}\left\{B_{H\mid T}(t)\overline{[\phi(g)-\mathrm{E}\{\phi(g)\}]}\right\}. \end{aligned}$$

However, it is clear that

$$\phi(g)-\mathrm{E}\{\phi(g)\} = \left[\frac{1}{\Gamma(3/2-H)}\right]^2 \int_0^T t^{2H-1}\left[\overline{\frac{d}{dt}\int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H} g'(\tau)d\tau}\right] d_t \int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H}\, dB_{H\mid T}(\tau)$$

$$= \langle \mathrm{B}_{H\mid T}, g\rangle_{\boldsymbol{H}(\mathrm{B}_{H\mid T})} \qquad \text{(by (3.3.16))}.$$

Hence,

$$\mathrm{E}\left\{B_{H\,|T}(t)\overline{[\phi(g)-\mathrm{E}\{\phi(g)\}]}\right\} = \mathrm{Cov}\left\{B_{H\,|T}(t), \langle \mathbf{B}_{H\,|T}, g \rangle_{\mathcal{H}(\mathbf{B}_{H|T})}\right\}$$

$$= g(t) \qquad \text{(by (2.1.1))},$$

and (3.4.9) follows. ■

For the sake of completeness, we give below a corollary to Theorem 3.4.1, which is simply a restatement in terms of the signal $s = m'$. The results of Theorem 3.4.2 are already in terms of m' so no restatement is given.

Corollary 3.4.3: Problem (3.4.1) is nonsingular if and only if, for almost all $t \in [0,T]$,

$$s(t) = \frac{1}{\Gamma(H-\frac{1}{2})} t^{H-\frac{1}{2}} \int_0^t (t-\tau)^{H-\frac{3}{2}} \tau^{\frac{1}{2}-H} \tilde{s}(\tau) d\tau, \tag{3.4.10}$$

where $\tilde{s} \in L^2([0,T])$, and, for almost all $t \in [0,T]$,

$$\tilde{s}(t) = t^{H-\frac{1}{2}} \frac{d}{dt} \left[\frac{1}{\Gamma(\frac{3}{2}-H)} \int_0^t (t-\tau)^{\frac{1}{2}-H} \tau^{\frac{1}{2}-H} s(\tau) d\tau \right]. \tag{3.4.11}$$

Proof: Problem (3.4.1) is nonsingular if and only if (3.4.2) is nonsingular, and (3.4.2) is nonsingular if and only if m satisfies (3.4.4). It follows from the proof of Theorem 3.3.4, that m satisfies (3.4.4) if and only if $s = m'$ satisfies (3.4.10). Equation (3.4.11) is simply a restatement of (3.4.5) in terms of s. ■

Remark: As a practical consideration, the representation of the likelihood ratio given by Equations (3.4.6) through (3.4.8) is probably rather difficult to implement. One would prefer to use a matched filter implementation, if possible. This is easily done if s is sufficiently well-behaved. For example, suppose

$$s(t) = V_H H(2H-1)\int_0^T h(\tau)\,|t-\tau|^{2H-2}d\tau,$$

where h is a continuous function on $[0,T]$. It then follows immediately that the likelihood ratio is given by (3.4.6) with

$$\langle \mathbf{Y}, m\rangle_{\mathbf{H}(\mathbf{Y})} = \int_0^T h(\tau)dY(\tau), \tag{3.4.12}$$

and

$$\langle m, m\rangle_{\mathbf{H}(\mathbf{Y})} = V_H H(2H-1)\int_0^T\int_0^T h(\tau)h(\sigma)\,|\tau-\sigma|^{2H-2}d\tau d\sigma. \tag{3.4.13}$$

Of course, in this case, even if **Y** is non-Gaussian (but with the same covariance function $K_{\mathbf{B}_{H|T}}$), the statistic $\langle \mathbf{Y}, m\rangle_{\mathbf{H}(\mathbf{Y})}$ given by (3.4.12) is the linear statistic with the highest signal-to-noise ratio.

3.5. Detection of Gaussian Signals in FGN

Let $\mathbf{X} \triangleq \{X(t);\ t \in [0,T]\}$ be a process observed on a compact interval $[0,T]$. In this section, we consider the following hypothesis testing problem:

H_0: **X** is Gaussian with mean zero and covariance function $K_{\mathbf{B}_{H|T}}$

versus (3.5.1)

H_1: **X** is Gaussian with mean function m_1 and continuous covariance function K_1.

This problem can be regarded as a generalization of the problem of detecting Gaussian signals in white Gaussian noise, which has been widely studied. We will see that the nonsingularity conditions for the detection Problem (3.5.1) are closely related to the following well-known result of Shepp [58].

Theorem 3.5.1: Let $\mathbf{X} \triangleq \{X(t);\ t \in [0,T]\}$ be a Gaussian process and let $K_0(t,s) = \min(t,s)$ for $t,s \in [0,T]$. (Recall that K_0 is the covariance function of a standard Brownian motion on $[0,T]$.) Consider the following hypothesis testing problem:

$$\begin{aligned} &H_0\text{: X has mean zero and covariance function } K_0 \\ \text{versus} & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (3.5.2) \\ &H_1\text{: X has mean function } m_1 \text{ and continuous covariance function } K_1. \end{aligned}$$

This problem is nonsingular if and only if the following three conditions are satisfied.

(i) There exists a function $\phi \in L^2([0,T] \times [0,T])$ such that

$$(K_1 - K_0)(t,s) = \int_0^t\int_0^s \phi(\tau,\sigma)\,d\sigma\,d\tau, \qquad \forall\ t,s \in [0,T].$$

(ii) If $\sigma(\Phi)$ represents the spectrum of the operator determined by the function ϕ, then $-1 \notin \sigma(\Phi)$.

(iii) There exists a function $\tilde{m}_1 \in L^2([0,T])$ such that

$$m_1(t) = \int_0^t \tilde{m}_1(\tau)\,d\tau, \qquad \forall\ t \in [0,T].$$

The function ϕ is unique and symmetric and is given by

$$\phi(t,s) = \frac{\partial^2}{\partial s\,\partial t}[(K_1 - K_0)(t,s)]$$

for almost all $(t,s) \in [0,T] \times [0,T]$. The function $\tilde{m}_1$ is unique and is given by

$$\tilde{m}_1(t) = \frac{d}{dt} m(t)$$

for almost all $t \in [0,T]$.

Shepp also gives formulas for the likelihood ratio in the event that Problem (3.5.2) is nonsingular. We will show that, if Problem (3.5.1) is nonsingular, it can be transformed into an equivalent problem of the form (3.5.2), in which case, the likelihood ratio can be found using one of Shepp's formulas. We begin by giving a characterization of the RKHS $\boldsymbol{H}(K_{B_{H|T}}) \otimes \boldsymbol{H}(K_{B_{H|T}})$.

Theorem 3.5.2: Let the set of functions $\{f_t;\, t \in [0,T]\}$ be defined by (3.3.10). Then $\boldsymbol{H}(K_{B_{H|T}}) \otimes \boldsymbol{H}(K_{B_{H|T}})$ consists of functions of the form

$$g(t,s) = \int_0^T\!\!\int_0^T f_t(\tau) f_s(\sigma) \overline{\tilde{g}(\tau,\sigma)}\, d\sigma\, d\tau, \qquad t,s \in [0,T], \tag{3.5.3}$$

where $\tilde{g} \in L^2([0,T] \times [0,T])$. Any such function g is twice differentiable almost everywhere in $[0,T] \times [0,T]$ and, for almost all $(t,s) \in [0,T] \times [0,T]$,

$$\overline{\tilde{g}(t,s)} = \left[\frac{1}{\Gamma(3/2-H)}\right]^2 t^{H-1/2} s^{H-1/2} \frac{\partial^2}{\partial s\, \partial t} \int_0^t\!\!\int_0^s \tau^{1/2-H} \sigma^{1/2-H} (t-\tau)^{1/2-H} (s-\sigma)^{1/2-H} \frac{\partial^2}{\partial\sigma \partial\tau} g(\tau,\sigma)\, d\sigma\, d\tau. \tag{3.5.4}$$

Proof: Recall from the proof of Theorem 3.3.4 that the functions $\{f_t;\, t \in [0,T]\}$ span $L^2([0,T])$ and

$$K_{B_{H|T}}(t,s) = \int_0^T f_t(u) f_s(u)\, du, \qquad \forall\, t,s \in [0,T].$$

It follows that the set of functions $\{f_{(t,s)};\ (t,s)\in[0,T]\times[0,T]\}$ given by

$$f_{(t,s)}(\tau,\sigma) \triangleq f_t(\tau)f_s(\sigma), \qquad \tau,\sigma\in[0,T],$$

spans $L^2([0,T]\times[0,T])$, and

$$K_{\mathbf{B}_{H|T}}(t_1,s_1)K_{\mathbf{B}_{H|T}}(t_2,s_2) = \int_0^T\!\!\int_0^T f_{(t_1,t_2)}(\tau,\sigma)f_{(s_1,s_2)}(\tau,\sigma)\,d\sigma\,d\tau, \qquad \forall\ t_i,s_i\in[0,T].$$

Since

$$K((t_1,t_2),(s_1,s_2)) = K_{\mathbf{B}_{H|T}}(t_1,s_1)K_{\mathbf{B}_{H|T}}(t_2,s_2), \qquad t_i,s_i\in[0,T],$$

is the reproducing kernel for $\boldsymbol{H}(K_{\mathbf{B}_{H|T}})\otimes\boldsymbol{H}(K_{\mathbf{B}_{H|T}})$, it follows from Theorem 2.1.1 that a function $g\in\boldsymbol{H}(K_{\mathbf{B}_{H|T}})\otimes\boldsymbol{H}(K_{\mathbf{B}_{H|T}})$ if and only if there exists a function $\tilde{g}\in L^2([0,T]\times[0,T])$ such that, for all $t,s\in[0,T]$,

$$\begin{aligned} g(t,s) &= \int_0^T\!\!\int_0^T f_{(t,s)}(\tau,\sigma)\overline{\tilde{g}(\tau,\sigma)}\,d\sigma\,d\tau \\ &= \int_0^T\!\!\int_0^T f_t(\tau)f_s(\sigma)\overline{\tilde{g}(\tau,\sigma)}\,d\sigma\,d\tau. \end{aligned}$$

This proves (3.5.3). To establish (3.5.4), notice that, if g satisfies (3.5.3), then

$$\begin{aligned} g(t,s) &= \int_0^T\!\!\int_0^T f_t(\tau)f_s(\sigma)\overline{\tilde{g}(\tau,\sigma)}\,d\sigma\,d\tau \\ &= \left[\frac{1}{\Gamma(H-\frac{1}{2})}\right]^2 \int_0^t\!\!\int_0^s \left[\tau^{\frac{1}{2}-H}\int_\tau^t u^{H-\frac{1}{2}}(u-\tau)^{H-\frac{3}{2}}du\right]\left[\sigma^{\frac{1}{2}-H}\int_\sigma^s v^{H-\frac{1}{2}}(v-\sigma)^{H-\frac{3}{2}}dv\right]\overline{\tilde{g}(\tau,\sigma)}\,d\sigma\,d\tau \\ &= \left[\frac{1}{\Gamma(H-\frac{1}{2})}\right]^2 \int_0^t\!\!\int_0^s u^{H-\frac{1}{2}}v^{H-\frac{1}{2}}\left[\int_0^u\!\!\int_0^v \tau^{\frac{1}{2}-H}\sigma^{\frac{1}{2}-H}(u-\tau)^{H-\frac{3}{2}}(v-\sigma)^{H-\frac{3}{2}}\overline{\tilde{g}(\tau,\sigma)}\,d\sigma\,d\tau\right]dv\,du. \end{aligned}$$

It follows that, for almost all $(t,s)\in[0,T]\times[0,T]$,

$$\frac{\partial^2}{\partial s\,\partial t} g(t,s) = \left[\frac{1}{\Gamma(H-\frac{1}{2})}\right]^2 t^{H-\frac{1}{2}} s^{H-\frac{1}{2}} \int_0^t\int_0^s \tau^{\frac{1}{2}-H}\,\sigma^{\frac{1}{2}-H}\,(t-\tau)^{H-\frac{3}{2}}(s-\sigma)^{H-\frac{3}{2}}\,\overline{g(\tau,\sigma)}\,d\sigma\,d\tau.$$

Equation (3.5.4) follows straightforwardly by using the above formula. ■

We can now prove the following theorem, which is analogous to Shepp's result for signal detection in white Gaussian noise.

Theorem 3.5.3: Problem (3.5.1) is nonsingular if and only if the following three conditions are satisfied.

(i) There exists a function $\phi \in L^2([0,T] \times [0,T])$ such that

$$(K_1 - K_{\mathbf{B}_{H|T}})(t,s) = \int_0^T\int_0^T f_t(\tau) f_s(\sigma)\phi(\tau,\sigma)\,d\sigma\,d\tau, \qquad \forall\; t,s \in [0,T]. \tag{3.5.5}$$

The function ϕ is real and symmetric, and, for almost all $(t,s) \in [0,T] \times [0,T]$,

$$\phi(t,s) = \left[\frac{1}{\Gamma(\frac{3}{2}-H)}\right]^2 t^{H-\frac{1}{2}} s^{H-\frac{1}{2}} \cdot \frac{\partial^2}{\partial s\,\partial t}\int_0^t\int_0^s \tau^{\frac{1}{2}-H}\,\sigma^{\frac{1}{2}-H}\,(t-\tau)^{\frac{1}{2}-H}\,(s-\sigma)^{\frac{1}{2}-H}\,\frac{\partial^2}{\partial\sigma\partial\tau}(K_1 - K_{\mathbf{B}_{H|T}})(\tau,\sigma)\,d\sigma\,d\tau. \tag{3.5.6}$$

(ii) If $\sigma(\Phi)$ represents the spectrum of the operator Φ determined by the function ϕ, then $\sigma(\Phi) \subset (-1,\infty)$.

(iii) There exists a function $\tilde{m}_1 \in L^2([0,T])$ such that

$$m_1(t) = \frac{1}{\Gamma(H-\frac{1}{2})}\int_0^t \tau^{\frac{1}{2}-H}\left[\int_\tau^t u^{H-\frac{1}{2}}(u-\tau)^{H-\frac{3}{2}}du\right]\tilde{m}_1(\tau)\,d\tau, \qquad \forall\; t \in [0,T]. \tag{3.5.7}$$

The function $\tilde{m}_1$ is real, and, for almost all $t \in [0,T]$,

$$\tilde{m}_1(t) = t^{H-\frac{1}{2}} \frac{d}{dt}\left[\frac{1}{\Gamma(\frac{3}{2}-H)}\int_0^t (t-\tau)^{\frac{1}{2}-H}\tau^{\frac{1}{2}-H} m'_1(\tau)d\tau\right]. \tag{3.5.8}$$

Proof: It follows immediately from Theorem 2.1.4 that Problem (3.5.1) is nonsingular if and only if

(i′) $(K_1 - K_{\mathrm{B}_{H|T}}) \in \boldsymbol{H}(K_{\mathrm{B}_{H|T}}) \otimes \boldsymbol{H}(K_{\mathrm{B}_{H|T}})$,

(ii′) There exist constants $0 < c < C < \infty$ such that $cK_{\mathrm{B}_{H|T}} \ll K_1 \ll CK_{\mathrm{B}_{H|T}}$, and

(iii′) $m_1 \in \boldsymbol{H}(K_{\mathrm{B}_{H|T}})$.

Clearly, by Theorem 3.5.2, condition (i) is equivalent to (i′), and it follows from Theorem 3.3.4 that condition (iii) is equivalent to (iii′). The theorem will be proven if we can show that condition (ii) is equivalent to (ii′). To this end, let ϕ be given by (3.5.6) and let the linear operator $\Phi: L^2([0,T]) \to L^2([0,T])$ be defined by

$$(\Phi g)(t) = \int_0^T \phi(t,\tau)g(\tau)d\tau, \qquad t \in [0,T],$$

where $g \in L^2([0,T])$. The operator Φ is then compact and self-adjoint with $\|\Phi\| \le \|\phi\|_2$. Further, $\sigma(\Phi) \subseteq [-\|\Phi\|, \|\Phi\|\,]$, and if $\lambda \in \sigma(\Phi)$, then $\lambda = 0$ or λ is an eigenvalue of Φ. Finally, if $\{\lambda_n\}_{n=1}^{\infty}$ is the set of eigenvalues of Φ and $\{e_n\}_{n=1}^{\infty}$ the associated set of orthonormal eigenfunctions, then zero is the only accumulation point of $\{\lambda_n\}_{n=1}^{\infty}$; $\{e_n\}_{n=1}^{\infty}$ spans $L^2([0,T])$; and

$$\phi(t,s) = \sum_{n=1}^{\infty} \lambda_n e_n(t)\overline{e_n(s)}, \qquad \forall\, t,s \in [0,T],$$

where the above convergence is in $L^2([0,T] \times [0,T])$. (These standard properties of L^2 integral operators are discussed, for example, in [41].) We are now ready to show that (ii) and (ii′) are equivalent.

(ii′) implies (ii): Suppose there exists $c < 0$ such that $(K_1 - cK_{\mathbf{B}_{H|T}})$ is nonnegative definite, and let λ be an eigenvalue of Φ with associated (unit) eigenfunction e. Since $\{f_t;\ t \in [0,T]\}$ spans $L^2([0,T])$, there exists a sequence of functions of the form

$$f_N(\tau) = \sum_{i=1}^{N} a_i f_{t_i}(\tau), \qquad \tau \in [0,T],$$

which converges to e in $L^2([0,T])$. Now, let δ be the Dirac delta function and recall that

$$K_{\mathbf{B}_{H|T}}(t,s) = \int_0^T f_t(\tau) f_s(\tau)\, d\tau = \int_0^T\!\!\int_0^T f_t(\tau) f_s(\sigma)\delta(\tau-\sigma)\, d\sigma\, d\tau.$$

Then,

$$\begin{aligned}
\lambda + (1-c) &= \int_0^T\!\!\int_0^T e(\sigma)\overline{e(\tau)}[\phi(\tau,\sigma) + (1-c)\delta(\tau-\sigma)]\, d\sigma\, d\tau \\
&= \lim_{N\to\infty} \int_0^T\!\!\int_0^T \left[\sum_{i=1}^{N} a_i f_{t_i}(\sigma)\right]\left[\overline{\sum_{j=1}^{N} a_j f_{t_j}(\tau)}\right][\phi(\tau,\sigma) + (1-c)\delta(\tau-\sigma)]\, d\sigma\, d\tau \\
&= \lim_{N\to\infty} \sum_{i,j=1}^{N} a_i \overline{a}_j (K_1 - cK_{\mathbf{B}_{H|T}})(t_i,t_j) \\
&\geq 0,
\end{aligned}$$

and it follows that $\lambda \geq c-1 > -1$. Hence, $\sigma(\Phi) \subseteq (-1, \|\Phi\|\,]$ and (ii′) implies (ii).

(ii) implies (ii′): Now suppose that $\sigma(\Phi) \subseteq (-1, \|\Phi\|\,]$. Let the sets $\{a_i\}_{i=1}^{N}$ and $\{t_i\}_{i=1}^{N}$ be arbitrary and define

$$g_N(\tau) \triangleq \sum_{i=1}^{N} a_i f_{t_i}(\tau), \qquad \tau \in [0,T].$$

Then

$$\sum_{i,j=1}^{N} a_i \overline{a}_j (K_1 - cK_{B_{H|T}})(t_i,t_j) = \int_0^T\!\!\int_0^T \left[\sum_{i=1}^{N} a_i f_{t_i}(\sigma)\right] \left[\overline{\sum_{j=1}^{N} a_j f_{t_j}(\tau)}\right] [\phi(\tau,\sigma) + (1-c)\delta(\tau-\sigma)]\, d\sigma\, d\tau$$

$$= \int_0^T\!\!\int_0^T g_N(\sigma)\overline{g_N(\tau)}[\phi(\tau,\sigma) + (1-c)\delta(\tau-\sigma)]\, d\sigma\, d\tau$$

$$= \sum_{n=1}^{\infty} [\lambda_n + (1-c)]\, |\langle g_N, e_n \rangle|^2$$

$$\geq [\inf\{\lambda_n\}_{n=1}^{\infty} + (1-c)]\, \|g_N\|_2^2$$

$$> 0 \qquad \text{if } c < 1 + \inf\{\lambda_n\}_{n=1}^{\infty}.$$

Since -1 cannot be an accumulation point of the set $\{\lambda_n\}_{n=1}^{\infty}$, we always have $\inf\{\lambda_n\}_{n=1}^{\infty} > -1$, and we can find $0 < c < 1 + \inf\{\lambda_n\}_{n=1}^{\infty}$ such that $(K_1 - cK_{B_{H|T}})$ is nonnegative definite. Similarly,

$$\sum_{i,j=1}^{N} a_i \overline{a}_j (CK_{B_{H|T}} - K_1)(t_i,t_j) = \int_0^T\!\!\int_0^T g_N(\sigma)\overline{g_N(\tau)}[(C-1)\delta(\tau-\sigma) - \phi(\tau,\sigma)]\, d\sigma\, d\tau$$

$$= \sum_{n=1}^{\infty} [(C-1) - \lambda_n]\, |\langle g_N, e_n \rangle|^2$$

$$\geq [(C-1) - \|\Phi\|\,]\, \|g_N\|_2^2$$

$$> 0 \qquad \text{if } C > 1 + \|\Phi\|.$$

Hence, we can find $0 < c < C < \infty$ such that $cK_{B_{H|T}} \ll K_1 \ll CK_{B_{H|T}}$. It follows that (ii) implies (ii′), and the theorem is proved. ■

We now wish to investigate the possibility of transforming Problem (3.5.1) into the more familiar form (3.5.2) considered by Shepp.

Theorem 3.5.4: If Problem (3.5.1) is nonsingular, the transformation

$$Z(t) \triangleq \frac{1}{\Gamma(3/2-H)} \int_0^t \tau^{H-1/2} d_\tau \int_0^\tau (\tau-u)^{1/2-H} u^{1/2-H} \, dX(u), \qquad t \in [0,T], \tag{3.5.9}$$

is well-defined under both hypotheses. The integral in (3.5.9) is to be interpreted as an integral with respect to the process $\mathbf{Y} \triangleq \{Y(t);\ t \in [0,T]\}$ given by

$$Y(t) \triangleq \frac{1}{\Gamma(3/2-H)} \int_0^t (t-\tau)^{1/2-H} \tau^{1/2-H} \, dX(\tau), \qquad t \in [0,T]. \tag{3.5.10}$$

The new observation process $\mathbf{Z} \triangleq \{Z(t);\ t \in [0,T]\}$ is Gaussian, and the transformed hypothesis testing problem becomes

H_0: $\mathbf{Z}$ has mean zero and covariance function $K_0(t,s) = \min(t,s)$.

versus (3.5.11)

H_1: $\mathbf{Z}$ has mean function m_Z and covariance function K_Z,

where

$$m_Z(t) = \int_0^t \tilde{m}_1(\tau) d\tau, \qquad t \in [0,T], \tag{3.5.12}$$

$$K_Z(t,s) = \int_0^t \int_0^s \phi(\tau,\sigma) d\sigma d\tau + \min(t,s), \qquad t,s \in [0,T], \tag{3.5.13}$$

and ϕ and $\tilde{m}_1$ are given by (3.5.6) and (3.5.8), respectively.

Proof: It follows from Theorem (3.3.5) that, under H_0, the transformations defined by (3.5.9) and (3.5.10) are well-defined, resulting in $\mathbf{Z}$ being a standard Brownian motion on $[0,T]$. Further, since Problem (3.5.1) is assumed to be nonsingular, it follows that (3.5.9) and (3.5.10) are also well-defined under H_1 (see, for example, [7], lemma 2.4). So we need only to establish that, under H_1, $\mathbf{Z}$ has a mean and covariance given by Equations (3.5.12) and

(3.5.13), respectively. To this end, let **Y** be as defined as in (3.5.10) and notice that, under H_1, **Y** has mean function

$$m_{\mathbf{Y}}(t) = \frac{1}{\Gamma(3/2-H)}\int_0^t (t-\tau)^{1/2-H}\tau^{1/2-H}m'(\tau)d\tau = \int_0^t \tau^{1/2-H}\tilde{m}_1(\tau)d\tau, \qquad t \in [0,T], \tag{3.5.14}$$

and covariance function

$$K_{\mathbf{Y}}(t,s) = \left[\frac{1}{\Gamma(3/2-H)}\right]^2 \int_0^t\int_0^s (t-\tau)^{1/2-H}\tau^{1/2-H}(s-\sigma)^{1/2-H}\sigma^{1/2-H}\frac{\partial^2}{\partial\sigma\partial\tau}K_1(\tau,\sigma)d\sigma d\tau, \qquad t,s \in [0,T].$$

Clearly,

$$K_1(t,s) = (K_1 - K_{\mathbf{B}_{H|T}})(t,s) + K_{\mathbf{B}_{H|T}}(t,s),$$

and it follows from the proof of Theorem 3.3.5 that

$$\frac{1}{2-2H}\min(t^{2-2H},s^{2-2H}) = \left[\frac{1}{\Gamma(3/2-H)}\right]^2 \int_0^t\int_0^s (t-\tau)^{1/2-H}\tau^{1/2-H}(s-\sigma)^{1/2-H}\sigma^{1/2-H}\frac{\partial^2}{\partial\sigma\partial\tau}K_{\mathbf{B}_{H|T}}(\tau,\sigma)d\sigma d\tau.$$

Thus, using these relationships and (3.5.6), we get straightforwardly that

$$K_{\mathbf{Y}}(t,s) = \int_0^t\int_0^s \tau^{1/2-H}\sigma^{1/2-H}\phi(\tau,\sigma)d\sigma d\tau + \frac{1}{2-2H}\min(t^{2-2H},s^{2-2H}), \qquad s,t \in [0,T]. \tag{3.5.15}$$

Hence, we have established that, under H_1, process **Y** has a mean and covariance given by (3.5.14) and (3.5.15), respectively. Now, writing **Z** as

$$Z(t) = \int_0^t \tau^{H-1/2}dY(\tau), \qquad t \in [0,T],$$

and proceeding in a manner similar to the above, it follows straightforwardly that, under H_1,

Z has a mean and covariance given by (3.5.12) and (3.5.13), respectively. This establishes the theorem. ■

We have shown that, if Problem (3.5.1) is nonsingular, (3.5.9) can be used to transform Problem (3.5.1) causally into the form (3.5.2), which corresponds to the problem of detecting a Gaussian signal in white Gaussian noise. Furthermore, since (3.5.9) is invertible under H_0 (via (3.3.11)), the two problems are equivalent. Since Shepp has given formulas for the likelihood ratio for Problem (3.5.2), we have a characterization for the optimal solution to Problem (3.5.1).

As a final consideration, we restrict our attention to a special case of Problem (3.5.1). In particular, we assume that, under H_1, **X** is Gaussian with $m_1 = 0$, and

$$K_1(t,s) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega}\overline{\frac{e^{i\omega s}-1}{i\omega}} S_1(\omega)d\omega, \qquad \forall\, t,s \in [0,T], \tag{3.5.16}$$

where S_1 is a spectral "density" satisfying

$$\liminf_{|\omega|\to\infty} |\omega|^n S_1(\omega) > 0, \tag{3.5.17}$$

and

$$\int_{-\infty}^{\infty} \frac{S_1(\omega)}{(1+\omega^2)^n} d\omega < \infty \tag{3.5.18}$$

for some integer n. The restriction that $m_1 = 0$ is not really necessary, merely convenient, while conditions (3.5.16) through (3.5.18) simply imply that, under H_1, the observed process has stationary increments with well-behaved spectral densities. (See [77] and [78] for a discussion of processes with stationary increments.) These conditions are not very restrictive, and when they hold, the observation process **X** can be defined as

$$X(t) = \int_0^t W(\tau)d\tau \triangleq \mathbf{W}(I_{[0,t]}), \qquad t \in [0,T],$$

where **W** is a generalized stationary Gaussian process.

If we let $S_{\mathbf{W}_H}$ be the spectral density of FGN, that is,

$$S_{\mathbf{W}_H}(\omega) = |\omega|^{1-2H}, \qquad 0 \neq \omega \in I\!R,$$

then we can rewrite Problem (3.5.1) in the following equivalent form:

$$\mathrm{H}_0\text{: } \mathbf{W} \text{ has spectral density } S_{\mathbf{W}_H}$$

versus (3.5.19)

$$\mathrm{H}_1\text{: } \mathbf{W} \text{ has spectral density } S_1.$$

With regard to this formulation, we have the following result.

Proposition 3.5.5: If there exits $\beta > 2H - ½$ such that

$$\lim_{|\omega| \to \infty} |\omega|^{\beta} [S_1(\omega) - S_{\mathbf{W}_H}(\omega)] = 0,$$

then (3.5.19) is nonsingular. If

$$\liminf_{|\omega| \to \infty} |\omega|^{2H-½} [S_1(\omega) - S_{\mathbf{W}_H}(\omega)] > 0,$$

then (3.5.19) is singular.

The proof of Proposition 3.5.5 is essentially due to Rozanov. For **W** a stationary L^2-process with integrable spectral densities, the result is proven in [56]. The extension to the case considered here follows easily from Lemmas 8 and 12 in [57].

For a wide class of covariance functions K_1, this result gives simple nonsingularity conditions for Problem (3.5.1). For example, if, under H_1,

$$X(t) = B_{H' \mid T}(t) + B_{H \mid T}(t), \qquad t \in [0,T],$$

where $\mathbf{B}_{H' \mid T} \triangleq \{B_{H' \mid T}(t);\, t \in [0,T]\}$ is another FBM independent of $\mathbf{B}_{H \mid T}$, then Problem (3.5.1) is nonsingular if and only if $H' > H + 1/4$. Heuristically, Proposition 3.5.5 implies that a Gaussian signal produces a nonsingular shift in additive FGN when the signal's spectrum has $o(|f|^{1/2-2H})$ tails.

3.6. Performance Characteristics

In this section, we investigate some aspects of detector performance in the presence of additive fractional Gaussian noise. In particular, we derive the signal-to-noise ratios of the optimal detectors for a class of baseband signals. We also derive channel-mismatch degradation factors for two members of this class.

Consider the following normalized version of Problem (3.4.1):

$$H_0:\ dY(t) = \frac{\sigma}{\sqrt{V_H}} dB_{H \mid T}(t), \qquad t \in [0,T],$$

versus (3.6.1)

$$H_1:\ dY(t) = s(t)dt + \frac{\sigma}{\sqrt{V_H}} dB_{H \mid T}(t), \qquad t \in [0,T],$$

where the normalizing factor $\sigma/\sqrt{V_H}$ reflects the assumption that the noise has "power" σ^2 (i.e., $\mathrm{Var}\{Y(1)\} = \sigma^2$). It follows from Corollary 3.4.3 that Problem (3.6.1) will be nonsingular if and only if, for almost all $t \in [0,T]$,

$$s(t) = \frac{\sigma}{\sqrt{V_H}\,\Gamma(H-\frac{1}{2})} t^{H-\frac{1}{2}} \int_0^t (t-\tau)^{H-\frac{3}{2}} \tau^{\frac{1}{2}-H} \tilde{s}(\tau) d\tau, \tag{3.6.2}$$

where $\tilde{s} \in L^2([0,T])$, and, for almost all $t \in [0,T]$,

$$\tilde{s}(t) = \frac{\sqrt{V_H}}{\sigma\Gamma(\frac{3}{2}-H)} t^{H-\frac{1}{2}} \frac{d}{dt} \left[\int_0^t (t-\tau)^{\frac{1}{2}-H} \tau^{\frac{1}{2}-H} s(\tau) d\tau \right]. \tag{3.6.3}$$

Assuming nonsingularity, the log-likelihood ratio for (3.6.1) is then given by (cf. (3.4.6))

$$\log \mathrm{L}(\mathrm{Y}) = \langle \mathrm{Y}, m \rangle_{\mathcal{H}(\mathrm{Y})} - \tfrac{1}{2} \langle m, m \rangle_{\mathcal{H}(\mathrm{Y})}, \tag{3.6.4}$$

where

$$m(t) \triangleq \int_0^t s(\tau) d\tau, \tag{3.6.5}$$

$$\langle m, m \rangle_{\mathcal{H}(\mathrm{Y})} = \int_0^T [\tilde{s}(\tau)]^2 d\tau, \tag{3.6.6}$$

and

$$\langle \mathrm{Y}, m \rangle_{\mathcal{H}(\mathrm{Y})} = \frac{\sqrt{V_H}}{\sigma\Gamma(\frac{3}{2}-H)} \int_0^T \tilde{s}(t) t^{H-\frac{1}{2}} d_t \int_0^t (t-\tau)^{\frac{1}{2}-H} \tau^{\frac{1}{2}-H} dY(\tau). \tag{3.6.7}$$

It follows from (2.1.2) and (2.1.3) that log L(Y) is distributed in the following manner under the two hypotheses:

$$\mathrm{H}_0\colon\ \log \mathrm{L}(\mathrm{Y}) \sim \mathrm{N}\left\{ -\tfrac{1}{2} \langle m, m \rangle_{\mathcal{H}(\mathrm{Y})}, \langle m, m \rangle_{\mathcal{H}(\mathrm{Y})} \right\},$$

$$\mathrm{H}_1\colon\ \log \mathrm{L}(\mathrm{Y}) \sim \mathrm{N}\left\{ \tfrac{1}{2} \langle m, m \rangle_{\mathcal{H}(\mathrm{Y})}, \langle m, m \rangle_{\mathcal{H}(\mathrm{Y})} \right\}.$$

Assuming equal prior probabilities on H_0 and H_1, the minimum-probability-of-error detection rule takes the form (see [54])

$$\delta(\mathrm{Y}) = \begin{cases} 1 & \text{if } \log \mathrm{L(Y)} \geq 0 \\ 0 & \text{if } \log \mathrm{L(Y)} < 0 \end{cases},$$

and the minimum probability of error is $\mathrm{P}_e^* = \Phi(-\tfrac{1}{2}\sqrt{\langle m, m\rangle_{\mathcal{H}(\mathrm{Y})}})$, where Φ is the standard normal distribution function. Note that P_e^* is a decreasing function of $\langle m, m\rangle_{\mathcal{H}(\mathrm{Y})}$, which represents the signal-to-noise ratio (SNR) of the optimal detector for Problem (3.6.1).

Now, suppose that the signal s is of the form

$$s(t) = \frac{C\sqrt{2\alpha+1}}{T^\alpha} t^\alpha, \qquad t \in [0,T], \tag{3.6.8}$$

where C is any real constant and $\alpha \geq 0$. All of these signals have average power C^2 and satisfy (3.6.2) with $\bar{s}$ given by

$$\begin{aligned}
\bar{s}(t) &= \frac{C\sqrt{V_H(2\alpha+1)}}{\sigma T^\alpha \Gamma(3/2-H)} t^{H-1/2} \frac{d}{dt}\left[\int_0^t (t-\tau)^{1/2-H} \tau^{1/2-H} s(\tau)\, d\tau\right] \\
&= \frac{C\sqrt{V_H(2\alpha+1)}}{\sigma T^\alpha \Gamma(3/2-H)} t^{H-1/2} \frac{d}{dt}\left[\int_0^t (t-\tau)^{1/2-H} \tau^{\alpha+1/2-H}\, d\tau\right] \\
&= \frac{C\sqrt{V_H(2\alpha+1)}}{\sigma T^\alpha \Gamma(3/2-H)} t^{H-1/2} \frac{d}{dt}\left[t^{\alpha+2-2H}\int_0^1 (1-\tau)^{1/2-H} \tau^{\alpha+1/2-H}\, d\tau\right] \\
&= \frac{C(\alpha+2-2H)\sqrt{V_H(2\alpha+1)}}{\sigma T^\alpha \Gamma(3/2-H)} t^{\alpha+1/2-H} \int_0^1 (1-\tau)^{1/2-H} \tau^{\alpha+1/2-H}\, d\tau \\
&= \frac{C(\alpha+2-2H)\sqrt{V_H(2\alpha+1)}}{\sigma T^\alpha \Gamma(3/2-H)} t^{\alpha+1/2-H} \frac{\Gamma(3/2-H)\Gamma(\alpha+3/2-H)}{\Gamma(\alpha+3-2H)} \\
&= \frac{C\,\Gamma(\alpha+3/2-H)\sqrt{V_H(2\alpha+1)}}{\sigma T^\alpha \Gamma(\alpha+2-2H)} t^{\alpha+1/2-H}.
\end{aligned}$$

Hence, the SNR of the optimal detector is

$$\langle m, m \rangle_{\mathcal{H}(Y)} = \int_0^T [\tilde{s}(\tau)]^2 d\tau$$

$$= \frac{C^2 V_H (2\alpha+1)}{\sigma^2(2\alpha+2-2H)} \left[\frac{\Gamma(\alpha+3/2-H)}{\Gamma(\alpha+2-2H)} \right]^2 T^{2-2H} . \tag{3.6.9}$$

For $H = 1/2$ expression (3.6.9) reduces to

$$\langle m, m \rangle_{\mathcal{H}(Y)} = \frac{C^2}{\sigma^2} T ,$$

as one would expect, since $H = 1/2$ corresponds to white noise. In this case, the optimal SNR grows linearly with respect to the length of the observation interval. However, for any H in the range $1/2 < H < 1$, $\langle m, m \rangle_{\mathcal{H}(Y)}$ grows *sublinearly* with respect to T. In fact, for $H \approx 1$, the performance of the optimal detector is virtually independent of the length of the observation interval. This is in spite of the fact that the signal energy grows linearly with T and is in marked contrast to the asymptotically linear growth that one would expect if the noise process had a rational spectrum.

It is interesting to note that a constant signal ($\alpha = 0$) is essentially the only periodic signal that displays this behavior. More precisely, if s is a continuously differentiable periodic signal, the optimal SNR will grow asymptotically at a linear rate for all H in the range $1/2 \leq H < 1$. The proof of this fact is straightforward, although somewhat messy, and is omitted.

For fixed T, the behavior of the optimal SNR, as a function of α and H, is pretty much as one would expect, that is, increasing in both α and H. In fact, (3.6.9) implies that, for fixed T,

$$\lim_{H \to 1} \langle m, m \rangle_{\mathcal{H}(Y)} = \infty, \qquad \forall\, \alpha > 0,$$

and

$$\lim_{\alpha \to \infty} \langle m, m \rangle_{\mathcal{H}(Y)} = \infty, \qquad \forall\, H > ½.$$

The behavior when $\alpha = 0$ is interesting. In this case, expression (3.6.9) reduces to

$$\langle m, m \rangle_{\mathcal{H}(Y)} = \frac{C^2 \Gamma(3/2 - H)}{\sigma^2 2H\, \Gamma(H + ½) \Gamma(3 - 2H)} T^{2-2H},$$

and, as $H \to 1$, we get

$$\langle m, m \rangle_{\mathcal{H}(Y)} \to \frac{C^2}{\sigma^2}.$$

An optimal SNR equal to C^2/σ^2 corresponds to the case in which the sample paths of the noise process are constant with probability one. In fact, if we were to extend the definition of FGN (with power σ^2) to include the limiting case, $H = 1$, we would want just such a process since

$$\lim_{H \to 1} \frac{\sigma^2}{V_H} K_{W_H}(\tau) = \lim_{H \to 1} \sigma^2 H (2H - 1) |\tau|^{2H-2} = \sigma^2, \qquad \forall\, \tau \in \mathbb{R}.$$

Hence, heuristically, FGN can be regarded as going from a completely *unpredictable* process at $H = ½$ to a completely *predictable* process at $H = 1$.

Now suppose that s is of the form (3.6.8) but that H is incorrectly identified; that is, suppose that log L(Y) is computed using (3.6.3) - (3.6.7) on the basis of $H' \neq H$. Then log L(Y) (which, in this case, is not the true log-likelihood ratio) is distributed as follows:

$$H_0\text{:}\ \log L(Y) \sim N(-\tfrac{1}{2}\mu,\ \nu^2),$$

$$H_1\text{:}\ \log L(Y) \sim N(\ \tfrac{1}{2}\mu,\ \nu^2),$$

where (using (3.6.9))

$$\mu = E_1\left\{\frac{\sqrt{V_{H'}}}{\sigma\Gamma(\tfrac{3}{2}-H')}\int_0^T \tilde{s}(t)t^{H'-\frac{1}{2}}d_t\int_0^t (t-\tau)^{\frac{1}{2}-H'}\tau^{\frac{1}{2}-H'}dY(\tau)\right\}$$

$$= \frac{\sqrt{V_{H'}}}{\sigma\Gamma(\tfrac{3}{2}-H')}\int_0^T \tilde{s}(t)t^{H'-\frac{1}{2}}d_t\int_0^t (t-\tau)^{\frac{1}{2}-H'}\tau^{\frac{1}{2}-H'}s(\tau)d\tau \tag{3.6.10}$$

$$= \int_0^T [\tilde{s}(\tau)]^2 d\tau$$

$$= \frac{C^2V_{H'}(2\alpha+1)}{\sigma^2(2\alpha+2-2H)}\left[\frac{\Gamma(\alpha+\tfrac{3}{2}-H')}{\Gamma(\alpha+2-2H')}\right]^2 T^{2-2H'},$$

and

$$\nu^2 = \mathrm{Var}\left\{\frac{\sqrt{V_{H'}}}{\sigma\Gamma(3/2-H')}\int_0^T \bar{s}(t)t^{H'-1/2}d_t\int_0^t (t-\tau)^{1/2-H'}\tau^{1/2-H'}dY(\tau)\right\}$$

$$= \mathrm{Var}\left\{\frac{CV_{H'}\Gamma(\alpha+3/2-H')\sqrt{2\alpha+1}}{\sigma\sqrt{V_H}T^{\alpha}\Gamma(\alpha+2-2H')\Gamma(3/2-H')}\int_0^T t^{\alpha}d_t\int_0^t (t-\tau)^{1/2-H'}\tau^{1/2-H'}dB_{H\,|T}(\tau)\right\}$$

$$= \mathrm{Var}\left\{\frac{CV_{H'}\Gamma(\alpha+3/2-H')\sqrt{2\alpha+1}}{\sigma\sqrt{V_H}T^{\alpha}\Gamma(\alpha+2-2H')\Gamma(3/2-H')}\left[T^{\alpha}\int_0^T (t-\tau)^{1/2-H'}\tau^{1/2-H'}dB_{H\,|T}(\tau)\right.\right.$$

$$\left.\left. - \alpha\int_0^T t^{\alpha-1}\left[\int_0^t (t-\tau)^{1/2-H'}\tau^{1/2-H'}dB_{H\,|T}(\tau)\right]dt\right]\right\}. \qquad (3.6.11)$$

We consider two special cases.

Case 1: Constant Signal ($\alpha = 0$)

If $\alpha = 0$, expression (3.6.10) can be rewritten as

$$\mu = \frac{C^2V_{H'}\beta(3/2-H', 3/2-H')}{\sigma^2\Gamma(2-2H')}T^{2-2H'},$$

where β represents the beta function, and (3.6.11) reduces to

$$\nu^2 = \left[\frac{CV_{H'}}{\sigma\Gamma(2-2H')}\right]^2 H(2H-1)\int_0^T\int_0^T (T-\tau)^{1/2-H'}\tau^{1/2-H'}(T-\upsilon)^{1/2-H'}\upsilon^{1/2-H'}|\tau-\upsilon|^{2H-2}d\tau d\upsilon$$

$$= \left[\frac{CV_{H'}}{\sigma\Gamma(2-2H')}\right]^2 H(2H-1)T^{2+2H-4H'}\int_0^1\int_0^1 (1-\tau)^{1/2-H'}\tau^{1/2-H'}(1-\upsilon)^{1/2-H'}\upsilon^{1/2-H'}|\tau-\upsilon|^{2H-2}d\tau d\upsilon$$

$$= \left[\frac{CV_{H'}}{\sigma\Gamma(2-2H')}\right]^2 H(2H-1)T^{2+2H-4H'}e(H,H'),$$

where

$$e(H,H') \triangleq \int_0^1\!\!\int_0^1 (1-\tau)^{\frac{1}{2}-H'}\tau^{\frac{1}{2}-H'}(1-\upsilon)^{\frac{1}{2}-H'}\upsilon^{\frac{1}{2}-H'}\,|\tau-\upsilon|^{2H-2}\,d\tau\,d\upsilon.$$

Hence, the probability of error if $\alpha = 0$ and $H' \neq H$ is

$$\begin{aligned} P_e &= \Phi\left[-\frac{\mu}{2v}\right] \\ &= \Phi\left[\frac{-C\,\beta(\frac{3}{2}-H',\frac{3}{2}-H')}{2\sigma\sqrt{H(2H-1)e(H,H')}}T^{1-H}\right]. \end{aligned}$$

Naturally, letting $H' = H$ gives us the minimum probability of error; that is,

$$\begin{aligned} P_e^* &= \Phi(-\tfrac{1}{2}\sqrt{\langle m, m\rangle_{\mathcal{H}(Y)}}) \\ &= \Phi\left[\frac{-C\,\beta(\frac{3}{2}-H,\frac{3}{2}-H)}{2\sigma\sqrt{H(2H-1)e(H,H)}}T^{1-H}\right]. \end{aligned}$$

It follows that

$$\begin{aligned} P_e &= \Phi\left[\frac{-C\,\beta(\frac{3}{2}-H',\frac{3}{2}-H')}{2\sigma\sqrt{H(2H-1)e(H,H')}}T^{1-H}\right]. \\ &= \Phi\left[\frac{-C\,\beta(\frac{3}{2}-H,\frac{3}{2}-H)}{2\sigma\sqrt{H(2H-1)e(H,H)}}T^{1-H}\sqrt{\frac{e(H,H)}{e(H,H')}}\,\frac{\beta(\frac{3}{2}-H',\frac{3}{2}-H')}{\beta(\frac{3}{2}-H,\frac{3}{2}-H)}\right] \\ &= \Phi\left[-\tfrac{1}{2}\sqrt{\langle m, m\rangle_{\mathcal{H}(Y)}\frac{e(H,H)}{e(H,H')}\left[\frac{\beta(\frac{3}{2}-H',\frac{3}{2}-H')}{\beta(\frac{3}{2}-H,\frac{3}{2}-H)}\right]^2}\,\right]. \end{aligned}$$

Hence, the *degradation factor* $\delta(H,H')$, which represents the loss in SNR due to channel mismatch, is given by

$$\delta(H,H') = \frac{e(H,H)}{e(H,H')}\left[\frac{\beta(\frac{3}{2}-H',\frac{3}{2}-H')}{\beta(\frac{3}{2}-H,\frac{3}{2}-H)}\right]^2.$$

The above derivation was carried out assuming that both H and H' were contained in the interval $(\frac{1}{2},1)$. However, it is straightforward to show that all of the above holds if

$H' = ½$ and $½ < H < 1$; that is, if a nonwhite channel is assumed to be white. If the actual channel is white, however, the expression for v^2 takes a different form. It is easy to see that, if $H = ½$ and $½ \le H' < 1$, we get

$$v^2 = \left[\frac{CV_{H'}}{\sigma\Gamma(2-2H)}\right]^2 T^{3-4H'}\int_0^1 (1-\tau)^{1-2H'}\tau^{1-2H'}d\tau$$

$$= \left[\frac{CV_{H'}}{\sigma\Gamma(2-2H)}\right]^2 T^{3-4H'}\beta(2-2H',\, 2-2H').$$

In this case, the degradation factor is given by

$$\delta(½,H') = \frac{[\,\beta(3/2-H',\, 3/2-H')\,]^2}{\beta(2-2H',\, 2-2H')}.$$

We have evaluated these degradation factors numerically for various values of H and H'. The results are presented in Table 3.6.1. As this table shows, if the channel is actually white ($H = ½$), the detector performance is very sensitive to mismatch. In fact, as $H' \to 1$, we have $\delta(½,H') \to 0$. On the other hand, Table 3.6.1 shows that, if the actual channel is nonwhite ($H > ½$), the detector performance is fairly *insensitive* to mismatch. In fact, if $H \approx 1$, there is virtually *no* loss in performance regardless of what value of H' is used to design the detector.

Clearly then, in the constant signal case, there is little to be gained, and potentially much to be lost, by using a detector designed to match an assumed value of $H > ½$. The robust strategy is to use a simple integrate-and-dump detector. This behavior is apparently related to the fact that, when using constant signals on a channel for which $H > ½$, both the signal and noise power are concentrated at the origin in the frequency domain. As we show below,

Table 3.6.1. Degradation $\delta(H, H')$ Due to Mismatch (Constant Signal)

H \ H'	0.500	0.600	0.700	0.800	0.900	0.950	0.975	0.999
0.500	1.0000	0.9910	0.9528	0.8534	0.6140	0.3823	0.2160	0.0098
0.600	0.9950	1.0000	0.9962	0.9622	0.8716	0.7850	0.7247	0.6533
0.700	0.9864	0.9960	1.0000	0.9934	0.9645	0.9341	0.9124	0.8862
0.800	0.9815	0.9902	0.9970	1.0000	0.9950	0.9870	0.9809	0.9734
0.900	0.9847	0.9898	0.9946	0.9984	1.0000	0.9994	0.9986	0.9972
0.950	0.9906	0.9932	0.9960	0.9982	0.9998	1.0000	1.0000	0.9996
0.975	0.9948	0.9962	0.9976	0.9988	0.9998	1.0000	1.0000	1.0000
0.999	0.9998	0.9998	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000

when using nonconstant signals, there is much more to be gained by properly identifying the self-similarity parameter H.

Case 2: Ramp Signal ($\alpha = 1$)

For $\alpha = 1$, (3.6.10) and (3.6.11) become

$$\mu = \frac{3C^2 V_{H'}}{\sigma^2(4-2H')} \left[\frac{\Gamma(5/2-H')}{\Gamma(3-2H')}\right]^2 T^{2-2H'},$$

and

$$v^2 = \left[\frac{\sqrt{3}CV_{H'}\Gamma(5/2-H')}{\sigma T\,\Gamma(3-2H')\Gamma(3/2-H')}\right]^2 H(2H-1)T^{4+2H-4H'}$$

$$\cdot \left\{\int_0^1\!\!\int_0^1 (1-\tau)^{1/2-H'}\tau^{1/2-H'}(1-\upsilon)^{1/2-H'}\upsilon^{1/2-H'}\,|\tau-\upsilon|^{2H-2}\right.$$

$$\left.\cdot\left[-\frac{2(1-\upsilon)}{3/2-H'}+\frac{(1-\tau)(1-\upsilon)}{(3/2-H')^2}\right]d\tau\,d\upsilon\right\}$$

$$= \left[\frac{\sqrt{3}CV_{H'}\Gamma(5/2-H')}{\sigma\Gamma(3-2H')\Gamma(3/2-H')}\right]^2 H(2H-1)T^{2+2H-4H'}e(H,H'),$$

where

$$e(H,H') \triangleq \int_0^1\!\!\int_0^1 (1-\tau)^{1/2-H'}\tau^{1/2-H'}(1-\upsilon)^{1/2-H'}\upsilon^{1/2-H'}\,|\tau-\upsilon|^{2H-2}$$

$$\cdot\left[1-\frac{2(1-\upsilon)}{3/2-H'}+\frac{(1-\tau)(1-\upsilon)}{(3/2-H')^2}\right]d\tau\,d\upsilon.$$

Proceeding as before, we find that, for $1/2 < H < 1$ and $1/2 \le H' < 1$,

$$\delta(H,H') = \frac{\Gamma(5/2-H')\Gamma(3/2-H')\Gamma(3-2H)(4-2H)}{\Gamma(5/2-H)\Gamma(3/2-H)\Gamma(3-2H')(4-2H')}\sqrt{\frac{e(H,H)}{e(H,H')}}.$$

Further, for $H = 1/2$ and $1/2 \le H' < 1$, it is straightforward to show that

$$\delta(1/2,H') = \frac{\Gamma(5/2-H')\Gamma(3/2-H')}{\Gamma(3-2H')(4-2H')}\sqrt{\frac{3}{e(1/2,H')}},$$

where

$$e(1/2,H') \triangleq \beta(2-2H',\,2-2H') - \frac{2}{3/2-H'}\beta(3-2H',\,2-2H') + \frac{1}{(3/2-H')^2}\beta(4-2H',\,2-2H').$$

Again, we have evaluated these degradation factors for several values of H and H'. The results are presented in Table 3.6.2. As this table shows, detector performance continues to be quite sensitive to channel mismatch when $H = \frac{1}{2}$. However, with this signal structure, detector performance is also quite sensitive to mismatch when $H > \frac{1}{2}$. This would seem to indicate that, for general nonconstant signals, significant performance improvements can be expected due to proper identification of the parameter H.

Table 3.6.2. Degradation $\delta(H, H')$ Due to Mismatch (Ramp Signal)

H \ H'	0.500	0.600	0.700	0.800	0.900	0.950	0.975	0.999
0.500	1.0000	0.9712	0.8505	0.5964	0.2615	0.1129	0.0514	0.0019
0.600	0.9783	1.0000	0.9600	0.7896	0.4791	0.3109	0.2357	0.1722
0.700	0.9084	0.9675	1.0000	0.9330	0.6745	0.4908	0.3996	0.3179
0.800	0.7691	0.8488	0.9415	1.0000	0.8701	0.6912	0.5845	0.4808
0.900	0.5081	0.5796	0.6901	0.8588	1.0000	0.9216	0.8188	0.6912
0.950	0.2980	0.3453	0.4272	0.5864	0.8906	1.0000	0.9561	0.8308
0.975	0.1623	0.1896	0.2392	0.3476	0.6498	0.9281	1.0000	0.9136
0.999	0.0071	0.0083	0.0107	0.0165	0.0419	0.1214	0.3341	1.0000

3.7. Conclusion

In this chapter, we have considered the problem of detecting signals in the presence of additive fractional Gaussian noise. We have applied results from the theory of reproducing kernel Hilbert spaces to give necessary and sufficient conditions for the problem to be non-singular and to develop whitening filters. For the case of a stationary stochastic signal, we have interpreted these results in the frequency domain. In the case when the signal is deterministic, we have characterized the optimal detector in terms of the likelihood ratio. Finally, we have studied some aspects of detector performance on FGN channels.

An interesting problem that we have not addressed in this context is that of sequence detection in FGN. It is obvious that, in the presence of such strongly dependent noise, the use of any one-shot strategy (optimal or not) to detect a sequence of signals will lead to a sequence of strongly dependent detection errors. The study of this phenomenon and its consequences is an interesting topic for further investigation.

CHAPTER 4

AN RKHS APPROACH TO ROBUST DETECTION AND ESTIMATION

4.1. Introduction

In this chapter, we consider the problems of L^2 estimation and signal detection in the presence of uncertainty regarding the statistical structure of the random processes involved. The classical approach to designing estimation and detection procedures relies upon exact knowledge of this statistical structure, and procedures designed to be optimal for a particular nominal structure sometimes perform very poorly if the actual statistical structure varies even slightly from the nominal. The minimax approach to designing robust procedures that display some amount of tolerance to variations in the actual statistical structure of the problem has been studied in recent years by many authors. The work presented here is most closely related to and motivated by the results presented in [52], [53], [66], and [70]. Other related results can be found, for example, in [9], [10], [24], [39], [40], [65], [68], and [69].

We formulate and analyze the robust estimation and detection problems in the context of reproducing kernel Hilbert space theory. Although the relationship between classical detection/estimation and RKHS theory is well-known (see, for example, [22], [23], and [48]), this theory has not been applied previously to the study of robust estimation and detection. By using an RKHS approach, we are able to generalize the notion of a linear filter and to give necessary and sufficient conditions for such a filter to be robust in the minimax sense for the general L^2-estimation problem in which there is uncertainty in both the covariance structure of the observed process $\mathbf{X} \triangleq \{X(t)\}$ and the cross-covariance structure of $\mathbf{X}$ and Z, the variable to be estimated. We show that, under mild regularity conditions, the robust filter can be found by solving a related minimization problem. We also give conditions sufficient to insure that the robust filter exists. In particular, we show that, if the covariance structure of the observed process is assumed to be known, so that the only uncertainty is in the cross-covariance structure of X and Z, then a robust filter will always exist and can be found by solving a straightforward minimization problem.

A somewhat surprising consequence of this analysis is the striking similarity between these results for the general robust L^2-estimation problem and results given by Poor in [53] relating to robust matched filtering. Reformulating the robust matched filtering problem in an RKHS context allows us to extend Poor's results and clearly reveals the underlying similarity between the robust estimation and matched filtering problems. In fact, the structures of minimax solutions to the two problems are seen to be virtually identical.

As a final application of the RKHS approach to robustness, we consider the problem of robust quadratic detection of a Gaussian signal in the presence of Gaussian noise, in which the deflection ratio is used as a performance criterion. We show that this problem also can be formulated in an RKHS context, and, when the structure of the noise covariance is assumed to be known, is exactly analogous to the robust matched filtering problem. If the covariance structure of the noise is also unknown, the robust quadratic detection problem can be embedded in a larger problem, which is again analogous to the robust matched filtering problem. A robust filter for this larger problem will then possess desirable robustness properties when applied to the quadratic detection problem.

This chapter is organized as follows. In Section 4.2, we present some definitions and notation that will be used throughout the chapter. In Section 4.3, we discuss the general robust L^2-estimation problem, and in Section 4.4 the robust matched filtering problem. Robust quadratic detection is discussed in Section 4.5. Section 4.6 presents some concluding remarks.

4.2. Preliminaries

Throughout this chapter, $\mathbf{X} \triangleq \{X(t);\ t \in I\}$ will represent an observed process defined on an arbitrary index set I. All random variables will be assumed real-valued with mean zero and finite variance unless otherwise specified. The extension of all results to the case of complex-valued random variables is straightforward.

Suppose that $\mathbb{K}$ is a class of covariance functions defined on an index set I. We define $\mathbf{H}(\mathbb{K})$ as the set of all functions belonging to $\mathbf{H}(K)$ for some $K \in \mathbb{K}$; that is,

$$\mathbf{H}(\mathbb{K}) = \bigcup_{K \in \mathbb{K}} \mathbf{H}(K).$$

Note that $\mathbf{H}(\mathbb{K})$ is generally *not* an RKHS, nor is it necessarily closed under addition. However, if $\mathbb{K}$ is a convex set of covariance functions, then $\mathbf{H}(\mathbb{K})$ is convex. This follows from the fact that if $s_0 \in \mathbf{H}(K_0)$ and $s_1 \in \mathbf{H}(K_1)$, then $(1-\alpha)s_0 + \alpha s_1 \in H((1-\alpha)K_0 + \alpha K_1)$ (see [1], §I.6).

We define a *finite filter on* $\mathbb{K}$ as a pair $\tilde{\mathbf{h}} = (\{h_i\}_{i=1}^n, \{t_i\}_{i=1}^n)$, where n is a positive integer, $\{h_i\}_{i=1}^n \subset \mathbb{R}$, and $\{t_i\}_{i=1}^n \subseteq I$. For each $K \in \mathbb{K}$, the function $\tilde{\mathbf{h}}K \in \mathbf{H}(K)$ is defined by

$$\tilde{\mathbf{h}}K(\cdot) \triangleq \sum_{i=1}^{n} h_i K(\cdot, t_i).$$

Similarly, we define a *filter on* $\mathbb{K}$ as a sequence $\mathbf{h} = \{\tilde{\mathbf{h}}_N\}_{N=1}^{\infty}$ of finite filters such that, for every $K \in \mathbb{K}$, the sequence $\{\bar{\mathbf{h}}_N K\}_{N=1}^{\infty} \subset \mathbf{H}(K)$ converges in $\mathbf{H}(K)$. For each $K \in \mathbb{K}$, the function $\mathbf{h}K \in \mathbf{H}(K)$ is then defined as

$$\mathbf{h}K(\cdot) \triangleq \lim_{N\to\infty} \bar{\mathbf{h}}_N K(\cdot) = \lim_{N\to\infty} \sum_{i=1}^{n(N)} h_i(N) K(\cdot, t_i(N)),$$

where the limit is taken in $\mathbf{H}(K)$ (and consequently pointwise, as well). In order to simplify notation somewhat, we will regard $\bigcup_{N=1}^{\infty} \{t_i(N)\}_{i=1}^{n(N)}$ as a single sequence $\{t_i\}_{i=1}^{\infty}$ and write

$$\mathbf{h}K(\cdot) = \lim_{N\to\infty} \sum_{i=1}^{\infty} h_i(N) K(\cdot, t_i),$$

where it is understood that, for each N, the sequence $\{h_i(N)\}_{i=1}^{\infty}$ has only finitely many nonzero elements corresponding to the appropriate values of the sequence $\{t_i\}_{i=1}^{\infty}$.

For any class $\mathbb{K}$ of covariance functions on I, we denote by $\mathbb{F}(\mathbb{K})$ the class of all filters on $\mathbb{K}$, and we note that the class of finite filters can be regarded as a subset of $\mathbb{F}(\mathbb{K})$. To denote the set of filters defined on the singleton class $\mathbb{K} = \{K\}$, we simply write $\mathbb{F}(K)$; thus, $\mathbb{F}(\mathbb{K}) = \bigcap_{K \in \mathbb{K}} \mathbb{F}(K)$. For any $\mathbf{h} \in \mathbb{F}(\mathbb{K})$, we can define a function $\langle \cdot, \mathbf{h} \rangle : \mathbf{H}(\mathbb{K}) \to \mathbb{R}$ in the following manner. Let $K \in \mathbb{K}$ and $s \in \mathbf{H}(K)$. Then

$$\begin{aligned}
\langle s, \mathbf{h} \rangle &\triangleq \langle s, \mathbf{h}K \rangle_{\mathbf{H}(K)} \\
&= \langle s, \lim_{N\to\infty} \bar{\mathbf{h}}_N K \rangle_{\mathbf{H}(K)} \\
&= \lim_{N\to\infty} \langle s, \bar{\mathbf{h}}_N K \rangle_{\mathbf{H}(K)} \\
&= \lim_{N\to\infty} \langle s, \sum_{i=1}^{\infty} h_i(N) K(\cdot, t_i) \rangle_{\mathbf{H}(K)} \\
&= \lim_{N\to\infty} \sum_{i=1}^{\infty} h_i(N) s(t_i).
\end{aligned}$$

It is clear that $\langle \cdot, \mathbf{h} \rangle$ is well-defined on $\mathbf{H}(\mathbb{K})$ for any $\mathbf{h} \in \mathbb{F}(\mathbb{K})$. Further, for any $\mathbf{h} \in \mathbb{F}(\mathbb{K})$

and $K \in \mathbb{K}$, the restriction of $\langle \cdot, \mathbf{h} \rangle$ to $\mathbb{H}(K)$ defines a bounded linear functional on $\mathbb{H}(K)$. Also, for any $K \in \mathbb{K}$, we have

$$\begin{aligned}\langle \mathbf{h}K, \mathbf{h} \rangle &= \langle \mathbf{h}K, \mathbf{h}K \rangle_{\mathbb{H}(K)} \\ &= \lim_{N \to \infty} \langle \bar{\mathbf{h}}_N K, \bar{\mathbf{h}}_N K \rangle_{\mathbb{H}(K)} \\ &= \lim_{N \to \infty} \sum_{i,j=1}^{\infty} h_i(N) h_j(N) K(t_j, t_i).\end{aligned}$$

If the process X has covariance function $K_{\mathbf{X}} \in \mathbb{K}$, then, for each $\mathbf{h} \in \mathbb{F}(\mathbb{K})$, we can define $\mathbf{h}(\mathbf{X}) \in L^2(\mathbf{X})$ as

$$\begin{aligned}\mathbf{h}(\mathbf{X}) &\triangleq \lim_{N \to \infty} \bar{\mathbf{h}}_N(\mathbf{X}) \\ &\triangleq \lim_{N \to \infty} \sum_{i=1}^{\infty} h_i(N) X(t_i),\end{aligned} \tag{4.2.1}$$

where the limit is taken in the mean-square sense. The fact that this limit is well-defined follows from the congruence between $\mathbb{H}(K_{\mathbf{X}})$ and $L^2(\mathbf{X})$; that is,

$$\begin{aligned}\lim_{N,M \to \infty} \mathrm{E}\{[\bar{\mathbf{h}}_N(\mathbf{X}) - \bar{\mathbf{h}}_M(\mathbf{X})]^2\} &= \lim_{N,M \to \infty} \|\bar{\mathbf{h}}_N K_{\mathbf{X}} - \bar{\mathbf{h}}_M K_{\mathbf{X}}\|^2_{\mathbb{H}(K_{\mathbf{X}})} \\ &= 0.\end{aligned}$$

Note that (4.2.1) implies that $\mathbf{h}(\mathbf{X}) = \langle \mathbf{X}, \mathbf{h}K_{\mathbf{X}} \rangle_{\mathbb{H}(K_{\mathbf{X}})}$.

Finally, the term *signal* will usually refer to a function $s: I \to \mathbb{R}$, and *uncertainty class* will always mean a set of signal-covariance pairs (s, K). The term *admissible uncertainty class* refers to an uncertainty class $\mathbb{U}$ with the following additional properties:

(i) for all pairs (s, K), $s \in \mathbb{H}(K)$, and

(ii) $\mathbb{U}$ is convex; that is, if $(s_0, K_0) \in \mathbb{U}$ and $(s_1, K_1) \in \mathbb{U}$, then $((1-\alpha)s_0 + \alpha s_1, (1-\alpha)K_0 + \alpha K_1) \in \mathbb{U}$ for all $0 \le \alpha \le 1$.

Given any uncertainty class $\boldsymbol{U}$, we associate with it the class $\mathbb{K}(\boldsymbol{U})$ of covariance functions "contained in" $\boldsymbol{U}$; that is,

$$\mathbb{K}(\boldsymbol{U}) \triangleq \left\{K : (s, K) \in \boldsymbol{U} \text{ for some signal } s\right\}.$$

We are now ready to discuss the problem of robust L^2 estimation.

4.3. Robust L^2 Estimation

Let $\mathbf{X} \triangleq \{X(t);\ t \in I\}$ be an observed zero-mean process with covariance function $K_{\mathbf{X}}$, and let Z be an arbitrary zero-mean, L^2 random variable. If $\hat{Z}(\mathbf{X})$ is the projection of Z onto $L^2(\mathbf{X})$, then we know that

$$\mathrm{E}\{[Z - \hat{Z}(\mathbf{X})]^2\} = \min_{U \in L^2(X)} \mathrm{E}\{[Z - U]^2\}. \tag{4.3.1}$$

That is, $\hat{Z}(\mathbf{X}) \in L^2(\mathbf{X})$ is the linear estimate of Z that minimizes the mean squared error (MSE). Of course, $\hat{Z}(\mathbf{X})$ is also the unique element of $L^2(\mathbf{X})$ that satisfies the equation

$$\mathrm{E}\{X(t)Z\} = \mathrm{E}\{X(t)\hat{Z}(\mathbf{X})\}, \qquad \forall\, t \in I. \tag{4.3.2}$$

We can restate these relationships in RKHS terms by defining $s_Z : I \rightarrow \mathbb{R}$ as

$$s_Z(t) \triangleq \mathrm{E}\{X(t)Z\}, \qquad \forall\, t \in I.$$

It then follows from Equations (2.1.1) and (4.3.2) that $s_Z \in \boldsymbol{H}(K_{\mathbf{X}})$ and $\langle \mathbf{X}, s_Z \rangle_{\boldsymbol{H}(K_{\mathbf{X}})} = \hat{Z}(\mathbf{X})$. Hence, if $K_{\mathbf{X}}$ and s_Z are known, the linear estimate $\hat{Z}(\mathbf{X}) \in L^2(\mathbf{X})$ satisfying (4.3.1) can be determined as follows. Since $s_Z \in \boldsymbol{H}(K_{\mathbf{X}})$, there exists a filter $\mathbf{h} \in \mathbb{F}(K_{\mathbf{X}})$ satisfying $s_Z = \mathbf{h}K_{\mathbf{X}}$; that is, for all $t \in I$,

$$s_Z(t) = \mathbf{h}K_X(t) = \lim_{N\to\infty} \bar{\mathbf{h}}_N K_X(t) = \lim_{N\to\infty} \sum_{i=1}^{\infty} h_i(N) K_X(t,t_i). \tag{4.3.3}$$

It follows that $\hat{Z}(X) = \langle X, s_Z \rangle_{\boldsymbol{H}(K_X)}$ is given by

$$\hat{Z}(X) = \mathbf{h}(X) = \lim_{N\to\infty} \sum_{i=1}^{\infty} h_i(N) X(t_i).$$

Of course, Equation (4.3.3) is just a generalization of the well-known Wiener-Hopf equation.

The foregoing discussion is a restatement of classical L^2-estimation results in RKHS terms, but now consider the problem in which there is uncertainty in the structure of K_X or s_Z. In particular, suppose we know only that the pair (s_Z, K_X) belongs to some admissible uncertainty class $\boldsymbol{U}$. Recall that, for any pair $(s, K) \in \boldsymbol{U}$, we require that $s \in \boldsymbol{H}(K)$. Since we have already seen that, in general, we must have $s_Z \in \boldsymbol{H}(K_X)$, this restriction on the structure of $\boldsymbol{U}$ is quite natural for this problem. Now, since K_X and s_Z are no longer assumed to be known precisely, we cannot expect to find the estimate $\hat{Z}(X) \in L^2(X)$ that solves (4.3.1). Instead, we look for a filter $\mathbf{h}_R \in \mathbb{F}(\mathbb{K}(\boldsymbol{U}))$ that satisfies

$$\sup_{(s,K)\in \boldsymbol{U}} M(\mathbf{h}_R; (s,K)) = \inf_{\mathbf{h}\in \mathbb{F}(\mathbb{K}(\boldsymbol{U}))} \sup_{(s,K)\in \boldsymbol{U}} M(\mathbf{h}; (s,K)), \tag{4.3.4}$$

where $M(\mathbf{h}; (s,K))$ is the MSE associated with estimating Z by $\mathbf{h}(X)$ if $K_X = K$ and $s_Z = s$; that is,

$$
\begin{aligned}
M(\mathbf{h};\,(s,K)) &\triangleq \mathrm{E}\{[Z-\mathbf{h}(\mathrm{X})]^2\} \\
&= \mathrm{E}\{[Z - \lim_{N\to\infty}\sum_{i=1}^{\infty} h_i(N)X(t_i)]^2\} \\
&= \mathrm{E}\{Z^2\} - 2\left[\lim_{N\to\infty}\sum_{i=1}^{\infty} h_i(N)\mathrm{E}\{X(t_i)Z\}\right] + \lim_{N\to\infty}\sum_{i,j=1}^{\infty} h_i(N)h_j(N)\mathrm{E}\{X(t_i)X(t_j)\} \\
&= \mathrm{E}\{Z^2\} - 2\left[\lim_{N\to\infty}\sum_{i=1}^{\infty} h_i(N)s(t_i)\right] + \lim_{N\to\infty}\sum_{i,j=1}^{\infty} h_i(N)h_j(N)K(t_j,t_i) \\
&= \sigma_Z^2 - 2\langle s,\mathbf{h}\rangle + \langle \mathbf{h}K,\mathbf{h}\rangle.
\end{aligned}
$$

Notice that we are implicitly assuming that $\sigma_Z^2 \triangleq \mathrm{E}\{Z^2\}$ does not depend on the actual structure of K_X or s_Z - it is assumed fixed but not necessarily known. This is usually a reasonable requirement on the structure of the uncertainty in the problem (see [52] and [66] for some examples), and, in any case, it is equivalent to solving a "normalized" problem in which the performance criterion is MSE/σ_Z^2, and $\mathcal{U}$ consists of pairs (s,K) corresponding to the possible values of $(s_Z/\sigma_Z^2,\, K_X/\sigma_Z^2)$.

A filter $\mathbf{h}_R \in \mathbb{F}(\mathbb{K}(\mathcal{U}))$ satisfying (4.3.4) is called a *robust filter* for the *game* $(\mathbb{F}(\mathbb{K}(\mathcal{U})),\mathcal{U},M)$. In order to find a robust filter, we search for a *saddle point* for $(\mathbb{F}(\mathbb{K}(\mathcal{U})),\mathcal{U},M)$, which is defined to be a pair $(\mathbf{h}_R;\,(s_L,K_L)) \in \mathbb{F}(\mathbb{K}(\mathcal{U})) \times \mathcal{U}$ that satisfies

$$
M(\mathbf{h}_R;\,(s,K)) \le M(\mathbf{h}_R;\,(s_L,K_L)) \le M(\mathbf{h};\,(s_L,K_L)) \qquad \forall\ \mathbf{h} \in \mathbb{F}(\mathbb{K}(\mathcal{U})),\ (s,K) \in \mathcal{U}. \tag{4.3.5}
$$

If $(\mathbf{h}_R;\,(s_L,K_L))$ satisfies (4.3.5), then it is easy to see that (see [4], §2.3.1)

$$
\inf_{\mathbf{h}\in\mathbb{F}(\mathbb{K}(\mathcal{U}))}\ \sup_{(s,K)\in\mathcal{U}} M(\mathbf{h};\,(s,K)) = \sup_{(s,K)\in\mathcal{U}}\ \inf_{\mathbf{h}\in\mathbb{F}(\mathbb{K}(\mathcal{U}))} M(\mathbf{h};\,(s,K)) = M(\mathbf{h}_R;\,(s_L,K_L)).
$$

In particular, $\mathbf{h}_R$ satisfies (4.3.4). If we define

$$M^*(s,K) \triangleq \inf_{\mathbf{h}\in \mathbb{F}(\mathbb{K}(\boldsymbol{U}))} M(\mathbf{h};\,(s,K)),$$

then (4.3.5) clearly implies that

$$M^*(s_L,K_L) = M(\mathbf{h}_R;\,(s_L,K_L)) \geq M^*(s,K), \qquad \forall\,(s,K) \in \boldsymbol{U};$$

that is, (s_L,K_L) is *least favorable* for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})),\boldsymbol{U},M)$. We consider some of the properties of M and M^*.

Lemma 4.3.1: M is concave on $\boldsymbol{U}$ for fixed $\mathbf{h} \in \mathbb{F}(\mathbb{K}(\boldsymbol{U}))$.

Proof: Let $(s_i,K_i) \in \boldsymbol{U}$ for $i = 0, 1$ and $(s_\alpha,K_\alpha) = ((1-\alpha)s_0 + \alpha s_1,\,(1-\alpha)K_0 + \alpha K_1)$ for $0 \leq \alpha \leq 1$. Then, for fixed $\mathbf{h} \in \mathbb{F}(\mathbb{K}(\boldsymbol{U}))$,

$$\begin{aligned}
M(\mathbf{h};\,(s_\alpha,K_\alpha)) &= \sigma_Z^2 - 2\langle s_\alpha,\,\mathbf{h}\rangle + \langle \mathbf{h}K_\alpha,\,\mathbf{h}\rangle \\
&= \sigma_Z^2 - 2\left[\lim_{N\to\infty}\sum_{i=1}^{\infty} h_i(N)s_\alpha(t_i)\right] + \lim_{N\to\infty}\sum_{i,j=1}^{\infty} h_i(N)h_j(N)K_\alpha(t_j,t_i) \\
&= (1-\alpha)\left[\sigma_Z^2 - 2\left[\lim_{N\to\infty}\sum_{i=1}^{\infty} h_i(N)s_0(t_i)\right] + \lim_{N\to\infty}\sum_{i,j=1}^{\infty} h_i(N)h_j(N)K_0(t_j,t_i)\right] \\
&\quad + \alpha\left[\sigma_Z^2 - 2\left[\lim_{N\to\infty}\sum_{i=1}^{\infty} h_i(N)s_1(t_i)\right] + \lim_{N\to\infty}\sum_{i,j=1}^{\infty} h_i(N)h_j(N)K_1(t_j,t_i)\right] \\
&= (1-\alpha)\left[\sigma_Z^2 - 2\langle s_0,\,\mathbf{h}_0\rangle + \langle \mathbf{h}K_0,\,\mathbf{h}\rangle\right] \\
&\quad + \alpha\left[\sigma_Z^2 - 2\langle s_1,\,\mathbf{h}\rangle + \langle \mathbf{h}K_1,\,\mathbf{h}\rangle\right] \\
&= (1-\alpha)M(\mathbf{h};\,(s_0,K_0)) + \alpha M(\mathbf{h};\,(s_1,K_1)). \quad \blacksquare
\end{aligned}$$

Lemma 4.3.2: M^* is concave on $\boldsymbol{U}$.

Proof: Let (s_α,K_α) be as before. Then

$$
\begin{aligned}
M^*(s_\alpha, K_\alpha) &= \inf_{\mathbf{h}\in \mathbb{F}(\mathbb{K}(\mathbb{U}))} M(\mathbf{h};\, (s_\alpha, K_\alpha)) \\
&\geq \inf_{\mathbf{h}\in \mathbb{F}(\mathbb{K}(\mathbb{U}))} \left[(1-\alpha) M(\mathbf{h};\, (s_0, K_0)) + \alpha M(\mathbf{h};\, (s_1, k_1))\right] \\
&\geq (1-\alpha)\left[\inf_{\mathbf{h}\in \mathbb{F}(\mathbb{K}(\mathbb{U}))} M(\mathbf{h};\, (s_0, K_0))\right] + \alpha\left[\inf_{\mathbf{h}\in \mathbb{F}(\mathbb{K}(\mathbb{U}))} M(\mathbf{h};\, (s_1, K_1))\right] \\
&= (1-\alpha) M^*(s_0, K_0) + \alpha M^*(s_1, K_1). \quad \blacksquare
\end{aligned}
$$

Lemma 4.3.3: For any $(s_0, K_0) \in \mathbb{U}$,

$$
M^*(s_0, K_0) = \sigma_Z^2 - \langle s_0,\, s_0\rangle_{\mathbb{H}(K_0)}.
$$

Further, for any $\mathbf{h} \in \mathbb{F}(\mathbb{K}(\mathbb{U}))$, we have

$$
M(\mathbf{h};\, (s_0, K_0)) = M^*(s_0, K_0)
$$

if and only if $\mathbf{h}K_0 = s_0$.

Proof: Since $s_0 \in \mathbb{H}(K_0)$, there exists $\mathbf{h}_0 \in \mathbb{F}(K_0))$ such that $\mathbf{h}_0 K_0 = s_0$. For any $\mathbf{h} \in \mathbb{F}(\mathbb{K}(\mathbb{U})) \subset \mathbb{F}(K_0)$, we have

$$
\begin{aligned}
M(\mathbf{h};\, (s_0, K_0)) &= \sigma_Z^2 - 2\langle s_0,\, \mathbf{h}\rangle + \langle \mathbf{h}K_0,\, \mathbf{h}\rangle \\
&= \sigma_Z^2 - 2\langle \mathbf{h}_0 K_0,\, \mathbf{h}\rangle + \langle \mathbf{h}K_0,\, \mathbf{h}\rangle \\
&= \sigma_Z^2 - 2\langle \mathbf{h}_0 K_0,\, \mathbf{h}K_0\rangle_{\mathbb{H}(K_0)} + \langle \mathbf{h}K_0,\, \mathbf{h}K_0\rangle_{\mathbb{H}(K_0)} \\
&= \sigma_Z^2 + \langle (\mathbf{h}-\mathbf{h}_0)K_0,\, (\mathbf{h}-\mathbf{h}_0)K_0\rangle_{\mathbb{H}(K_0)} - \langle \mathbf{h}_0 K_0,\, \mathbf{h}_0 K_0\rangle_{\mathbb{H}(K_0)} \\
&= \sigma_Z^2 + \langle (\mathbf{h}-\mathbf{h}_0)K_0,\, (\mathbf{h}-\mathbf{h}_0)K_0\rangle_{\mathbb{H}(K_0)} - \langle s_0,\, s_0\rangle_{\mathbb{H}(K_0)}.
\end{aligned}
$$

Hence,

$$
M(\mathbf{h};\, (s_0, K_0)) \geq \sigma_Z^2 - \langle s_0,\, s_0\rangle_{\mathbb{H}(K_0)},
$$

with equality if and only if $\mathbf{h}K_0 = \mathbf{h}_0 K_0 = s_0$. Further, since $\mathbf{h}_0 \in \mathbb{F}(K_0)$, there exists a sequence $\{\bar{\mathbf{h}}_N^0\}_{N=1}^\infty \subset \mathbb{F}(\mathbb{K}(\mathbb{U}))$ of finite filters such that

$$\lim_{N\to\infty} \bar{\mathbf{h}}_N^0 K_0 = \mathbf{h}_0 K_0 = s_0,$$

and it follows that

$$\begin{aligned}
\lim_{N\to\infty} M(\bar{\mathbf{h}}_N^0;\, (s_0,K_0)) &= \sigma_Z^2 - 2\left[\lim_{N\to\infty} \langle s_0,\, \bar{\mathbf{h}}_N^0\rangle\right] + \lim_{N\to\infty} \langle \bar{\mathbf{h}}_N^0 K_0,\, \bar{\mathbf{h}}_N^0\rangle \\
&= \sigma_Z^2 - 2\left[\lim_{N\to\infty} \langle s_0,\, \bar{\mathbf{h}}_N^0 K_0\rangle_{\mathbf{H}(K_0)}\right] + \lim_{N\to\infty} \langle \bar{\mathbf{h}}_N^0 K_0,\, \bar{\mathbf{h}}_N^0 K_0\rangle_{\mathbf{H}(K_0)} \\
&= \sigma_Z^2 - 2\langle s_0,\, s_0\rangle_{\mathbf{H}(K_0)} + \langle s_0,\, s_0\rangle_{\mathbf{H}(K_0)} \\
&= \sigma_Z^2 - \langle s_0,\, s_0\rangle_{\mathbf{H}(K_0)}.
\end{aligned}$$

This proves the lemma. ■

We can now give necessary and sufficient conditions for a pair $(\mathbf{h}_R;\, (s_L,K_L))$ to be a saddle point for $(\mathbb{F}(\mathbb{K}(\mathbb{U})),\mathbb{U},M)$. These conditions are very reminiscent of those given by Poor in [53], Lemma 1, relating to saddle point solutions for the robust matched filtering problem. The similarities between these two problems will be explored more fully in the next section.

Theorem 4.3.4: $(\mathbf{h}_R;\, (s_L,K_L)) \in \mathbb{F}(\mathbb{K}(\mathbb{U})) \times \mathbb{U}$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\mathbb{U})),\mathbb{U},M)$ if and only if $\mathbf{h}_R K_L = s_L$, and

$$2\langle s,\, \mathbf{h}_R\rangle - \langle s_L,\, \mathbf{h}_R\rangle - \langle \mathbf{h}_R K,\, \mathbf{h}_R\rangle \geq 0, \qquad \forall\, (s,K) \in \mathbb{U}. \tag{4.3.6}$$

Proof: Recall that $(\mathbf{h}_R;\, (s_L,K_L))$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\mathbb{U})),\mathbb{U},M)$ if and only if

$$M(\mathbf{h}_R;\, (s_L,K_L)) \leq M(\mathbf{h};\, (s_L,K_L)), \qquad \forall\, \mathbf{h} \in \mathbb{F}(\mathbb{K}(\mathbb{U})), \tag{4.3.7}$$

and

$$M(\mathbf{h}_R;\, (s,K)) \leq M(\mathbf{h}_R;\, (s_L,K_L)), \qquad \forall\, (s,K) \in \mathbb{U}. \tag{4.3.8}$$

It follows immediately from Lemma 4.3.3 that (4.3.7) is satisfied if and only if $\mathbf{h}_R K_L = s_L$, so the theorem will be proven if we can establish that, given $\mathbf{h}_R K_L = s_L$, (4.3.8) holds if and only if (4.3.6) holds. To this end, notice that, since M is concave on $\boldsymbol{U}$ for fixed $\mathbf{h} \in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}))$, (4.3.8) holds if and only if, for each $(s, K) \in \boldsymbol{U}$,

$$\lim_{\alpha \to 0} \frac{1}{\alpha} \Big[M(\mathbf{h}_R;\, (s_\alpha, K_\alpha)) - M(\mathbf{h}_R;\, (s_L, K_L)) \Big] \le 0,$$

where $(s_\alpha, K_\alpha) = ((1-\alpha)s_L + \alpha s,\, (1-\alpha)K_L + \alpha K)$. But if $\mathbf{h}_R K_L = s_L$, then

$$\langle s_L, s_L \rangle_{\boldsymbol{H}(K_L)} = \langle s_L, \mathbf{h}_R \rangle = \langle \mathbf{h}_R K_L, \mathbf{h}_R \rangle,$$

and, by Lemma 4.3.3,

$$M(\mathbf{h}_R;\, (s_L, K_L)) = \sigma_Z^2 - \langle s_L, s_L \rangle_{\boldsymbol{H}(K_L)}.$$

Therefore,

$$\begin{aligned}
\lim_{\alpha \to 0} \frac{1}{\alpha} \Big[M(\mathbf{h}_R;\, (s_\alpha, K_\alpha)) - M(\mathbf{h}_R;\, (s_L, K_L)) \Big] &= \lim_{\alpha \to 0} \frac{1}{\alpha} \Big[-2\alpha \langle s, \mathbf{h}_R \rangle - 2(1-\alpha) \langle s_L, \mathbf{h}_R \rangle \\
&\qquad + \langle \mathbf{h}_R K_\alpha, \mathbf{h}_R \rangle + \langle \mathbf{h}_R K_L, \mathbf{h}_R \rangle \Big] \\
&= \lim_{\alpha \to 0} \frac{1}{\alpha} \Big[-2\alpha \langle s, \mathbf{h}_R \rangle - 2(1-\alpha) \langle s_L, \mathbf{h}_R \rangle \\
&\qquad + \alpha \langle \mathbf{h}_R K, \mathbf{h}_R \rangle + (2-\alpha) \langle \mathbf{h}_R K_L, \mathbf{h}_R \rangle \Big] \\
&= \lim_{\alpha \to 0} \frac{1}{\alpha} \Big[-2\alpha \langle s, \mathbf{h}_R \rangle + \alpha \langle s_L, \mathbf{h}_R \rangle + \alpha \langle \mathbf{h}_R K, \mathbf{h}_R \rangle \Big] \\
&= -2 \langle s, \mathbf{h}_R \rangle + \langle s_L, \mathbf{h}_R \rangle + \langle \mathbf{h}_R K, \mathbf{h}_R \rangle.
\end{aligned}$$

Hence, if $\mathbf{h}_R K_L = s_L$, then (4.3.8) holds if and only if (4.3.6) holds, and the theorem follows. ■

Theorem 4.3.4 gives a complete characterization of saddle point solutions for $(\boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U})), \boldsymbol{U}, M)$. The next theorem and its corollary give conditions under which the existence of a saddle point is equivalent to the existence of a least favorable pair (s_L, K_L)

maximizing M^* on $\boldsymbol{U}$. Since M^* is concave on $\boldsymbol{U}$, the search for a least favorable pair often reduces to a straightforward convex programming problem. The proof of Theorem 4.3.5 is given in Appendix B.

Theorem 4.3.5: Let $(s_L, K_L) \in \boldsymbol{U}$ and let $\mathbf{h}_R \in \mathbb{F}(K_L)$ satisfy $\mathbf{h}_R K_L = s_L$. Suppose that K_L *dominates* $\mathbb{K}(\boldsymbol{U})$; that is, for all $K \in \mathbb{K}(\boldsymbol{U})$, there exists $C > 0$ (depending on K) such that $K \ll CK_L$ (i.e., $(CK_L - K)$ is nonnegative definite on I). Then $\mathbf{h}_R \in \mathbb{F}(\mathbb{K}(\boldsymbol{U}))$ and the condition

$$\langle s_L, s_L \rangle_{\boldsymbol{H}(K_L)} \leq \langle s, s \rangle_{\boldsymbol{H}(K)}, \qquad \forall\, (s, K) \in \boldsymbol{U}, \tag{4.3.9}$$

holds if and only if (4.3.6) holds; that is, if and only if

$$2\langle s, \mathbf{h}_R \rangle - \langle s_L, \mathbf{h}_R \rangle - \langle \mathbf{h}_R K, \mathbf{h}_R \rangle \geq 0, \qquad \forall\, (s, K) \in \boldsymbol{U}.$$

Corollary 4.3.6: Suppose that (s_L, K_L) is a least favorable pair for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})), \boldsymbol{U}, M)$ and that K_L dominates $\mathbb{K}(\boldsymbol{U})$. Let $\mathbf{h}_R \in \mathbb{F}(K_L)$ be such that $\mathbf{h}_R K_L = s_L$. Then $(\mathbf{h}_R; (s_L, K_L))$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})), \boldsymbol{U}, M)$.

Proof: If (s_L, K_L) is a least favorable pair, then

$$\langle s_L, s_L \rangle_{\boldsymbol{H}(K_L)} \leq \langle s, s \rangle_{\boldsymbol{H}(K)}, \qquad \forall\, (s, K) \in \boldsymbol{U}.$$

Since K_L dominates $\mathbb{K}(\boldsymbol{U})$, it follows from Theorem 4.3.5 that $\mathbf{h}_R \in \mathbb{F}(\mathbb{K}(\boldsymbol{U}))$ and

$$2\langle s, \mathbf{h}_R \rangle - \langle s_L, \mathbf{h}_R \rangle - \langle \mathbf{h}_R K, \mathbf{h}_R \rangle \geq 0, \qquad \forall\, (s, K) \in \boldsymbol{U}.$$

The result follows by applying Theorem 4.3.4. ■

Remark 4.3.7: Suppose that the index set I is itself a Hilbert space $\boldsymbol{H}_0$. Then the covariance functions $K \in \mathbb{K}(\boldsymbol{U})$ all correspond to bounded linear operators on $\boldsymbol{H}_0$, and a sufficient condition for K_L to dominate $\mathbb{K}(\mathbb{U})$ is that the operator $\mathrm{K_L}$ be invertible. To see this, let $K \in \mathbb{K}(\mathbb{U})$ with corresponding operator K and $f \in \mathbb{H}_0$. Then

$$\begin{aligned}
\langle f, (C\mathrm{K_L} - \mathrm{K})f \rangle_{\mathbb{H}_0} &= \langle f, \mathrm{K_L^{1/2}}(C\mathrm{I} - \mathrm{K_L^{-1/2}KK_L^{-1/2}})\mathrm{K_L^{1/2}}f \rangle_{\mathbb{H}_0} \\
&= \langle \mathrm{K_L^{1/2}}f, (C\mathrm{I} - \mathrm{K_L^{-1/2}KK_L^{-1/2}})\mathrm{K_L^{1/2}}f \rangle_{\mathbb{H}_0} \\
&= C\|\mathrm{K_L^{1/2}}f\|^2_{\mathbb{H}_0} - \langle \mathrm{K_L^{1/2}}f, (\mathrm{K_L^{-1/2}KK_L^{-1/2}})\mathrm{K_L^{1/2}}f \rangle_{\mathbb{H}_0} \\
&\geq \|\mathrm{K_L^{1/2}}f\|^2_{\mathbb{H}_0} [\, C - \|\mathrm{K_L^{-1/2}KK_L^{-1/2}}\| \,].
\end{aligned}$$

Clearly, if C is chosen to satisfy $C \geq \|\mathrm{K_L^{-1/2}KK_L^{-1/2}}\|$, then the operator $(C\mathrm{K_L} - \mathrm{K})$ is positive and $K \ll CK_L$.

It follows that, if (s_L, K_L) is a least favorable pair for $(\mathbb{F}(\mathbb{K}(\mathbb{U})), \mathbb{U}, M)$ and the operator $\mathrm{K_L}$ is invertible, the pair $(\mathrm{K_L^{-1}}s_L; (s_L, K_L))$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\mathbb{U})), \mathbb{U}, M)$. Note that, in this case, the filter $\mathbf{h}_R$ consists of a single element $h_R = \mathrm{K_L^{-1}}s_L \in \mathbb{H}_0$. The corresponding random variable is $\mathbf{h}_R(\mathrm{X}) = \mathrm{X}(h_R)$.

In the case when the covariance structure of the process X is assumed to be known, so that the only uncertainty is in the structure of the "signal" s_Z, the existence of a robust filter is guaranteed, as the following corollary shows.

Corollary 4.3.8: Suppose $\mathbb{U}$ has the form

$$\mathbb{U} = \{(s, K_0) : s \in S \subset \mathbb{H}(K_0),\ S \text{ convex}\}$$

for some fixed covariance function K_0. Then there exists a robust filter $\mathbf{h}_R$ for the game $(\mathbb{F}(K_0), \mathbb{U}, M)$. Further, $\mathbf{h}_R$ satisfies $\mathbf{h}_R K_0 = s_L$, where s_L is the unique element of $\overline{S}$ (the closure of S in $\mathbb{H}(K_0)$) with minimum norm; that is,

$$\langle s_L, s_L \rangle_{\boldsymbol{H}(K_0)} \leq \langle s, s \rangle_{\boldsymbol{H}(K_0)}, \qquad \forall\, s_L \in \overline{S}. \tag{4.3.10}$$

Proof: Let $\overline{\boldsymbol{U}} = \{(s, K_0): s \in \overline{S}\}$. Since $\overline{S}$ is a closed, convex subset of $\boldsymbol{H}(K_0)$, there exists a unique element $s_L \in \overline{S}$ satisfying (4.3.10) (see [31], §3.12). Since $s_L \in \boldsymbol{H}(K_0)$, we can find $\mathbf{h}_R$ satisfying $\mathbf{h}_R K_0 = s_L$, and, by Corollary 4.3.6, $(\mathbf{h}_R; (s_L, K_0))$ is a saddle point for $(\boldsymbol{F}(K_0), \overline{\boldsymbol{U}}, M)$. Since $M^*(s, K_0) = \sigma_Z^2 - \langle s, s \rangle_{\boldsymbol{H}(K_0)}$ is a continuous function on $\overline{S}$, it follows that

$$\sup_{(s,K_0) \in \boldsymbol{U}} M^*(s, K_0) = \sup_{(s,K_0) \in \overline{\boldsymbol{U}}} M^*(s, K_0).$$

Hence, following [70], §II.D, we have

$$\begin{aligned}
M(\mathbf{h}_R; (s_L, K_0)) &= \sup_{(s,K_0) \in \overline{\boldsymbol{U}}} M(\mathbf{h}_R; (s, K_0)) \\
&\geq \sup_{(s,K_0) \in \boldsymbol{U}} M(\mathbf{h}_R; (s, K_0)) \\
&\geq \inf_{\mathbf{h} \in \boldsymbol{F}(K_0)} \sup_{(s,K_0) \in \boldsymbol{U}} M(\mathbf{h}; (s, K_0)) \\
&\geq \sup_{(s,K_0) \in \boldsymbol{U}} \inf_{\mathbf{h} \in \boldsymbol{F}(K_0)} M(\mathbf{h}; (s, K_0)) \\
&= \sup_{(s,K_0) \in \boldsymbol{U}} M^*(s, K_0) \\
&= \sup_{(s,K_0) \in \overline{\boldsymbol{U}}} M^*(s, K_0) \\
&= \sup_{(s,K_0) \in \overline{\boldsymbol{U}}} \inf_{\mathbf{h} \in \boldsymbol{F}(K_0)} M(\mathbf{h}; (s, K_0)) \\
&= M(\mathbf{h}_R; (s_L, K_0)).
\end{aligned}$$

It follows that all of the inequalities above can be replaced with equalities, and we get

$$\sup_{(s,K_0) \in \boldsymbol{U}} M(\mathbf{h}_R; (s, K_0)) = \inf_{\mathbf{h} \in \boldsymbol{F}(K_0)} \sup_{(s,K_0) \in \boldsymbol{U}} M(\mathbf{h}; (s, K_0)),$$

so that $\mathbf{h}_R$ is a robust filter for $(\boldsymbol{F}(K_0), \boldsymbol{U}, M)$. ∎

If the covariance structure of **X** is not precisely known, considerably more structure is needed to guarantee the existence of a robust filter. The following corollary gives one set of conditions sufficient to guarantee the existence of a saddle point for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})),\boldsymbol{U},M)$.

Corollary 4.3.9: Suppose that I is a Hilbert space $\boldsymbol{H}_0$, and let us regard $\mathbb{K}(\boldsymbol{U})$ as a set of bounded, linear, self-adjoint, positive operators on $\boldsymbol{H}_0$. Suppose further that $\mathbb{K}(\boldsymbol{U}) \subset \mathrm{HS}(\boldsymbol{H}_0)$, the Hilbert space of Hilbert-Schmidt operators on $\boldsymbol{H}_0$ with the Hilbert-Schmidt norm $\|\cdot\|_{\mathrm{HS}(\boldsymbol{H}_0)}$. We can then regard $\boldsymbol{U}$ as a subset of $\boldsymbol{H}_0 \times \mathrm{HS}(\boldsymbol{H}_0)$, which becomes a Hilbert space under the norm

$$\|(\cdot,*)\|^2 = \|\cdot\|^2_{\boldsymbol{H}_0} + \|*\|^2_{\mathrm{HS}(\boldsymbol{H}_0)}.$$

If

(i) for every $\mathrm{K}_0, \mathrm{K}_1 \in \mathbb{K}(\boldsymbol{U})$ there exist constants $0 < c < C < \infty$ (depending on K_0 and K_1) such that $c\,\mathrm{K}_0 \ll \mathrm{K}_1 \ll C\,\mathrm{K}_0$ (i.e., the operators $(C\,\mathrm{K}_0 - \mathrm{K}_1)$ and $(\mathrm{K}_1 - c\,\mathrm{K}_0)$ are both positive), and

(ii) $\boldsymbol{U}$ is a closed, bounded subset of $\boldsymbol{H}_0 \times \mathrm{HS}(\boldsymbol{H}_0)$,

then there exists a saddle point for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})),\boldsymbol{U},M)$.

Proof: For the case under consideration $(I = \boldsymbol{H}_0)$, it is easy to see that a finite filter $\bar{\mathbf{h}}$ corresponds to a function $\bar{h} \in \boldsymbol{H}_0$, and, for any $(s,\mathrm{K}) \in \boldsymbol{U}$,

$$\langle s, \bar{\mathbf{h}} \rangle = \langle \bar{h}, s \rangle_{\boldsymbol{H}_0},$$

$$\langle \bar{\mathbf{h}}\mathrm{K}, \bar{\mathbf{h}} \rangle = \langle \bar{h}, \mathrm{K}\bar{h} \rangle_{\boldsymbol{H}_0},$$

and

$$M(\bar{\mathbf{h}};\,(s,\mathrm{K})) = \sigma_Z^2 - 2\langle \bar{h},\, s\rangle_{\boldsymbol{H}_0} + \langle \bar{h},\, \mathrm{K}\bar{h}\rangle_{\boldsymbol{H}_0}.$$

Clearly, for fixed $\bar{\mathbf{h}}$, the function $M(\bar{\mathbf{h}};\,(\cdot,\,*))$ can be extended to a continuous, concave function on all of $\boldsymbol{H}_0 \times \mathrm{HS}(\boldsymbol{H}_0)$. By considering the proof of Lemma 4.3.3, it is easy to see that, for all $(s,\mathrm{K}) \in \boldsymbol{U}$,

$$\begin{aligned} M^*(s,\mathrm{K}) &\triangleq \inf_{\mathbf{h}\in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}))} M(\mathbf{h};\,(s,\mathrm{K})) \\ &= \inf_{\bar{\mathbf{h}}\in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}))} M(\bar{\mathbf{h}};\,(s,\mathrm{K})). \end{aligned}$$

Hence, M^* also can be extended to all of $\boldsymbol{H}_0 \times \mathrm{HS}(\boldsymbol{H}_0)$ as the infimum of a class of continuous, concave functions. It follows that M^* (so extended) is concave (see the proof of Lemma 4.3.2) and upper semicontinuous (see [15], Proposition 7.11). By hypothesis (ii), $\boldsymbol{U}$ is a closed, bounded, convex subset of the Hilbert space $\boldsymbol{H}_0 \times \mathrm{HS}(\boldsymbol{H}_0)$, and it follows from [4], Theorem 2.1.2, that there exists $(s_{\mathrm{L}},\mathrm{K}_{\mathrm{L}}) \in \boldsymbol{U}$ such that

$$M^*(s_{\mathrm{L}},\mathrm{K}_{\mathrm{L}}) = \sup_{(s,\mathrm{K})\in \boldsymbol{U}} M^*(s,\mathrm{K});$$

that is, $(s_{\mathrm{L}},\mathrm{K}_{\mathrm{L}})$ is a least favorable pair for $(\boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U})),\boldsymbol{U},M)$. By hypothesis (i), K_{L} dominates $\boldsymbol{K}(\boldsymbol{U})$, and it follows from Corollary 4.3.6 that $(\mathbf{h}_{\mathrm{R}};\,(s_{\mathrm{L}},\mathrm{K}_{\mathrm{L}}))$ is a saddle point for $(\boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U})),\boldsymbol{U},M)$ for any $\mathbf{h}_{\mathrm{R}} \in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}))$ satisfying $\mathbf{h}_{\mathrm{R}}\mathrm{K}_{\mathrm{L}} = s_{\mathrm{L}}$. ■

We end this section with an example illustrating the material discussed above.

Example 4.3.10: Let $I = L^2([a,b])$, the space of square-integrable, real-valued functions on the interval $[a,b]$, where $-\infty < a < b < \infty$. Let k_0 be a known, real-valued, continuous covariance function on $[a,b]$ and let K_0 be the operator on $L^2([a,b])$ generated by k_0; that is,

$$(\mathrm{K}_0 f)(\cdot) \triangleq \int_a^b f(\tau) k_0(\cdot,\tau) d\tau, \qquad \forall f \in L^2([a,b]).$$

Assume that K_0 is strictly positive. It follows (see, for example, [41], §6.10) that K_0 is a compact operator on $L^2([a,b])$, and, if we let $\{e_n\}_{n=1}^{\infty}$ be the set of orthonormal eigenfunctions of K_0 with associated (positive) eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots$, then $\{e_n\}_{n=1}^{\infty}$ spans $L^2([a,b])$. If we regard $\boldsymbol{H}(\mathrm{K}_0)$ as a subset of $L^2([a,b])$, then it follows (see [48], §9) that a function $s \in L^2([a,b])$ is a member of $\boldsymbol{H}(\mathrm{K}_0)$ if and only if

$$\|s\|^2_{\boldsymbol{H}(\mathrm{K}_0)} \triangleq \sum_{n=1}^{\infty} \frac{s_n^2}{\lambda_n} < \infty, \tag{4.3.11}$$

where

$$s_n = \int_a^b s(\tau) e_n(\tau) d\tau.$$

Now, suppose that $\sigma_Z^2 > 0$, $s_0 \in \boldsymbol{H}(\mathrm{K}_0)$, and $\varepsilon > 0$ are given such that $\|s_0\|^2_{\boldsymbol{H}(\mathrm{K}_0)} < \sigma_Z^2$ and $\varepsilon < \sum_{n=1}^{\infty} (s_n^0)^2$, where

$$s_n^0 = \int_a^b s_0(\tau) e_n(\tau) d\tau.$$

Define the set $S \subset \boldsymbol{H}(\mathrm{K}_0)$ by

$$S \triangleq \left\{ s \in \boldsymbol{H}(\mathrm{K}_0) : \|s\|^2_{\boldsymbol{H}(\mathrm{K}_0)} \le \sigma_Z^2, \text{ and } \int_a^b [s(\tau) - s_0(\tau)]^2 d\tau \le \varepsilon \right\}$$
$$= \left\{ s \in L^2([a,b]) : \sum_{n=1}^{\infty} \frac{s_n^2}{\lambda_n} \le \sigma_Z^2, \text{ and } \sum_{n=1}^{\infty} (s_n - s_n^0)^2 \le \varepsilon \right\}. \quad (4.3.12)$$

We seek a saddle point for the game $(\boldsymbol{F}(\mathrm{K}_0), \boldsymbol{U}, \boldsymbol{M})$, where $\boldsymbol{U} = S \times \{\mathrm{K}_0\}$. We proceed by searching for the least favorable signal; that is, the signal $s_{\mathrm{L}} \in S$ that satisfies

$$\|s_{\mathrm{L}}\|^2_{\boldsymbol{H}(\mathrm{K}_0)} = \inf_{s \in S} \|s\|^2_{\boldsymbol{H}(\mathrm{K}_0)}. \quad (4.3.13)$$

It follows from (4.3.11) and (4.3.12) that solving (4.3.13) is equivalent to minimizing $\sum_{n=1}^{\infty} \frac{s_n^2}{\lambda_n}$ subject to the constraint $\sum_{n=1}^{\infty} (s_n - s_n^0)^2 \le \varepsilon$. A straightforward Lagrange multiplier calculation yields the candidate $s_{\mathrm{L}} \in L^2([a,b])$ defined by

$$s_{\mathrm{L}} = \sum_{n=1}^{\infty} s_n^{\mathrm{L}} e_n \triangleq \sum_{n=1}^{\infty} \left[\frac{\sigma \lambda_n s_n^0}{\sigma \lambda_n + 1} \right] e_n, \quad (4.3.14)$$

where $\sigma > 0$ is the unique positive solution to

$$\sum_{n=1}^{\infty} (s_n^0 - s_n^{\mathrm{L}})^2 = \sum_{n=1}^{\infty} \left[\frac{s_n^0}{\sigma \lambda_n + 1} \right]^2 = \varepsilon. \quad (4.3.15)$$

Notice that (4.3.14) and (4.3.15) imply that

$$\sum_{n=1}^{\infty} \frac{(s_n^{\mathrm{L}})^2}{\lambda_n} \le \lambda_1 \sum_{n=1}^{\infty} \left[\frac{s_n^{\mathrm{L}}}{\lambda_n} \right]^2 = \lambda_1 \sigma^2 \sum_{n=1}^{\infty} \left[\frac{s_n^0}{\sigma \lambda_n + 1} \right]^2 = \lambda_1 \sigma^2 \varepsilon < \infty, \quad (4.3.16)$$

so that $s_{\mathrm{L}} \in \boldsymbol{H}(\mathrm{K}_0)$.

To show that s_L is indeed the least favorable signal, we use the fact that s_L satisfies (4.3.13) if and only if (see [31], §3.12, or apply Theorem 4.3.5)

$$\langle s - s_L, s_L \rangle_{\mathcal{H}(K_0)} \geq 0, \qquad \forall\, s \in S.$$

It follows from (4.3.11), (4.3.14), and (4.3.15) that

$$\begin{aligned}\langle s - s_L, s_L \rangle_{\mathcal{H}(K_0)} &= \langle s - s_0, s_L \rangle_{\mathcal{H}(K_0)} + \langle s_0 - s_L, s_L \rangle_{\mathcal{H}(K_0)} \\ &= \langle s - s_0, s_L \rangle_{\mathcal{H}(K_0)} + \sum_{n=1}^{\infty} \frac{1}{\lambda_n} \left[\frac{s_n^0}{\sigma\lambda_n + 1} \right] \left[\frac{\sigma\lambda_n s_n^0}{\sigma\lambda_n + 1} \right] \\ &= \langle s - s_0, s_L \rangle_{\mathcal{H}(K_0)} + \sigma\varepsilon.\end{aligned}$$

Further, it follows from (4.3.11), (4.3.12), (4.3.16), and the Schwarz inequality that, for all $s \in S$,

$$\begin{aligned}|\langle s - s_0, s_L \rangle_{\mathcal{H}(K_0)}| &= \left| \sum_{n=1}^{\infty} \frac{(s_n - s_n^0) s_n^L}{\lambda_n} \right| \\ &\leq \left[\sum_{n=1}^{\infty} (s_n - s_n^0)^2 \right]^{1/2} \left[\sum_{n=1}^{\infty} \left[\frac{s_n^L}{\lambda_n} \right]^2 \right]^{1/2} \\ &\leq \sigma\varepsilon.\end{aligned}$$

Hence,

$$\langle s - s_L, s_L \rangle_{\mathcal{H}(K_0)} \geq 0, \qquad \forall\, s \in S,$$

and it follows that s_L is the least favorable signal.

Now, define $h_R \in L^2([a,b])$ by

$$h_{\mathrm{R}} = \sum_{n=1}^{\infty} \frac{s_n^{\mathrm{L}}}{\lambda_n} e_n . \tag{4.3.17}$$

Then,

$$\mathrm{K}_0 h_{\mathrm{R}} = \sum_{n=1}^{\infty} \frac{s_n^{\mathrm{L}}}{\lambda_n} \mathrm{K}_0 e_n = \sum_{n=1}^{\infty} s_n^{\mathrm{L}} e_n = s_{\mathrm{L}},$$

and it follows from Corollary 4.3.6 that $(h_{\mathrm{R}}; (s_{\mathrm{L}}, \mathrm{K}_0))$ is a saddle point for $(\mathbb{F}(\mathrm{K}_0), \mathbb{U}, \mathcal{M})$. Note that, here again, the filter $\mathbf{h}_{\mathrm{R}}$ consists of a single element h_{R}. The corresponding random variable is

$$\mathbf{h}_{\mathrm{R}}(\mathrm{X}) = \mathrm{X}(h_{\mathrm{R}}) = \int_a^b h_{\mathrm{R}}(\tau) X(\tau) d\tau .$$

We now turn to the problem of robust matched filtering. As we shall see, the structure of this problem is essentially the same as that of robust L^2 estimation.

4.4. Robust Matched Filtering

Let $\mathrm{X} \triangleq \{X(t);\ t \in I\}$ be an observed process. Consider the following simple hypothesis testing problem:

$$\begin{aligned} &\mathrm{H}_0\text{: X has mean zero and covariance function } K_{\mathrm{X}} \\ &\text{versus} \\ &\mathrm{H}_1\text{: X has mean function } m \text{ and covariance function } K_{\mathrm{X}}, \end{aligned} \tag{4.4.1}$$

where K_{X} is a known covariance function and $0 \neq m \in \mathbf{H}(K_{\mathrm{X}})$ is a known deterministic signal. This corresponds, of course, to the problem of detecting the signal m in the presence of additive, zero-mean noise with covariance function K_{X}. The condition $m \in \mathbf{H}(K_{\mathrm{X}})$

guarantees that Problem (4.4.1) is well-behaved and, when $\mathbf{X}$ is Gaussian, is related to the nonsingularity of (4.4.1), as discussed in Chapter 2. We shall have more to say about this condition at a later time.

Problem (4.4.1) is usually decided by using a linear detector; that is, a detector in which the test statistic $\mathbf{h}(\mathbf{X})$ is of the form

$$\mathbf{h}(\mathbf{X}) = \lim_{N \to \infty} \sum_{i=1}^{\infty} h_i(N) X(t_i),$$

where, for each N, the sequence $\{h_i(N)\}_{i=1}^{\infty} \subset \mathbb{R}$ has only finitely many nonzero elements, and convergence is the mean-square sense under both hypotheses. Clearly, given any such statistic, there is a corresponding filter $\mathbf{h} \in \mathbb{F}(K_{\mathbf{X}})$. Furthermore, since $m \in \mathbb{H}(K_{\mathbf{X}})$, it is easy to see that, given any $\mathbf{h} \in \mathbb{F}(K_{\mathbf{X}})$, the corresponding statistic $\mathbf{h}(\mathbf{X})$ is well-defined under both hypotheses.

The *signal-to-noise ratio* (SNR) for Problem (4.4.1) corresponding to the filter $\mathbf{h} \in \mathbb{F}(K_{\mathbf{X}})$ with associated test statistic $\mathbf{h}(\mathbf{X})$ is given by[1]

---

[1]Here and elsewhere, the quantity $\frac{0}{0}$ is defined to be 0.

$$\begin{aligned}
\Delta(\mathbf{h};\,(m,K_X)) &\triangleq \frac{\left[E_1\{\mathbf{h}(X)\}\right]^2}{\mathrm{Var}_0\{\mathbf{h}(X)\}} \\
&= \frac{\left[\lim_{N\to\infty}\sum_{i=1}^{\infty} h_i(N)\, m(t_i)\right]^2}{\lim_{N\to\infty}\sum_{i=1}^{\infty} h_i(N)\, h_j(N)\, K_X(t_j, t_i)} \\
&= \frac{\left[\langle m, \mathbf{h}\rangle\right]^2}{\langle \mathbf{h}K_X, \mathbf{h}\rangle} \\
&= \frac{\left[\langle m, \mathbf{h}K_X\rangle_{\mathcal{H}(K_X)}\right]^2}{\langle \mathbf{h}K_X, \mathbf{h}K_X\rangle_{\mathcal{H}(K_X)}}.
\end{aligned} \tag{4.4.2}$$

A *matched filter* for Problem (4.4.1) is any filter $\mathbf{h}^* \in \mathcal{F}(K_X)$ that satisfies

$$\Delta(\mathbf{h}^*;\,(m,K_X)) = \sup_{\mathbf{h}\in\mathcal{F}(K_X)} \Delta(\mathbf{h};\,(m,K_X)). \tag{4.4.3}$$

It follows immediately from (4.4.2) and the Schwarz inequality that $\mathbf{h}^*$ satisfies (4.4.3) if and only if $\mathbf{h}^* K_X = cm$, for some constant $c \neq 0$, and the maximum SNR for Problem (4.4.1) is given by

$$\Delta(\mathbf{h}^*;\,(m,K_X)) = \|m\|^2_{\mathcal{H}(K_X)}.$$

Now, suppose that the pair (m,K_X) is known only to belong to some admissible uncertainty class $\mathcal{U}$; that is, we now consider the composite hypothesis testing problem:

$$H_0\text{: } \mathbf{X} \text{ has mean zero and covariance function } K,$$

versus (4.4.4)

$$H_1\text{: } \mathbf{X} \text{ has mean function } s \text{ and covariance function } K,$$

where $(s, K) \in \boldsymbol{U}$.[2] In order to decide Problem (4.4.4), we are interested in using a linear detector incorporating a test statistic $\mathbf{h}_R(\mathbf{X})$ corresponding to a filter $\mathbf{h}_R \in \mathbb{F}(\mathbb{K}(\boldsymbol{U}))$ that satisfies

$$\inf_{(s,K)\in \boldsymbol{U}} \Delta(\mathbf{h}_R; (s,K)) = \sup_{\mathbf{h}\in \mathbb{F}(\mathbb{K}(\boldsymbol{U}))} \inf_{(s,K)\in \boldsymbol{U}} \Delta(\mathbf{h}; (s,K)). \tag{4.4.5}$$

We will refer to a filter satisfying (4.4.5) as a *robust matched filter* for the game $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})), \boldsymbol{U}, \Delta)$. In order to find a robust matched filter, we again search for a *saddle point* $(\mathbf{h}_R; (s_L, K_L)) \in \mathbb{F}(\mathbb{K}(\boldsymbol{U})) \times \boldsymbol{U}$ satisfying

$$\Delta(\mathbf{h}; (s_L, K_L)) \le \Delta(\mathbf{h}_R; (s_L, K_L)) \le \Delta(\mathbf{h}_R; (s,K)), \qquad \forall\ \mathbf{h} \in \mathbb{F}(\mathbb{K}(\boldsymbol{U})),\ (s,K) \in \boldsymbol{U}. \tag{4.4.6}$$

If we define

$$\Delta^*(s,K) \triangleq \sup_{\mathbf{h}\in \mathbb{F}(\mathbb{K}(\boldsymbol{U}))} \Delta(\mathbf{h}; (s,K)),$$

then, as before, (4.4.6) implies

$$\Delta^*(s_L, K_L) \le \Delta^*(s,K), \qquad \forall\ (s,K) \in \boldsymbol{U},$$

so that (s_L, K_L) is *least favorable* for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})), \boldsymbol{U}, \Delta)$. We have the following properties for Δ and Δ^*.

[2] We assume throughout this section that, for any pair $(s,K) \in \boldsymbol{U}$, $s \neq 0$.

Lemma 4.4.1: Δ is convex on $\boldsymbol{U}$ for fixed $\mathbf{h} \in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}))$.

Proof: Let $(s_i, K_i) \in \boldsymbol{U}$ for $i = 0, 1$, and define (s_α, K_α) in the usual manner for $0 \le \alpha \le 1$. If $\langle \mathbf{h}K_i, \mathbf{h} \rangle = 0$ for $i = 0$ or $i = 1$, then it is easy to see that

$$\Delta(\mathbf{h}; (s_\alpha, K_\alpha)) = (1-\alpha)\Delta(\mathbf{h}; (s_0, K_0)) + \alpha\Delta(\mathbf{h}; (s_1, K_1)).$$

If $\langle \mathbf{h}K_i, \mathbf{h} \rangle > 0$ for $i = 0$ and $i = 1$, then, letting

$$\beta = \frac{\alpha \langle \mathbf{h}K_1, \mathbf{h} \rangle}{\langle \mathbf{h}K_\alpha, \mathbf{h} \rangle},$$

we have $\beta \in [0,1]$ and

$$\begin{aligned}
\Delta(\mathbf{h}; (s_\alpha, K_\alpha)) &= \langle \mathbf{h}K_\alpha, \mathbf{h} \rangle \left[(1-\beta)\frac{\langle s_0, \mathbf{h} \rangle}{\langle \mathbf{h}K_0, \mathbf{h} \rangle} + \beta \frac{\langle s_1, \mathbf{h} \rangle}{\langle \mathbf{h}K_1, \mathbf{h} \rangle} \right]^2 \\
&\le \langle \mathbf{h}K_\alpha, \mathbf{h} \rangle \left[(1-\beta) \left[\frac{\langle s_0, \mathbf{h} \rangle}{\langle \mathbf{h}K_0, \mathbf{h} \rangle} \right]^2 + \beta \left[\frac{\langle s_1, \mathbf{h} \rangle}{\langle \mathbf{h}K_1, \mathbf{h} \rangle} \right]^2 \right] \\
&= (1-\alpha)\Delta(\mathbf{h}; (s_0, K_0)) + \alpha\Delta(\mathbf{h}; (s_1, K_1)).
\end{aligned}$$

This proves the lemma. ■

Lemma 4.4.2: Δ^* is convex on $\boldsymbol{U}$.

Proof: See proof of Lemma 4.3.2.

Lemma 4.4.3: For any $(s_0, K_0) \in \boldsymbol{U}$,

$$\Delta^*(s_0, K_0) = \langle s_0, s_0 \rangle_{\boldsymbol{H}(K_0)}.$$

Further, for any $\mathbf{h} \in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}))$, we have

$$\Delta(\mathbf{h}; (s_0, K_0)) = \Delta^*(s_0, K_0)$$

if and only if $\mathbf{h}K_0 = cs_0$ for some $c \neq 0$.

Proof: By the Schwarz inequality, we have

$$\Delta(\mathbf{h};\,(s_0,K_0)) = \frac{\left[\langle s_0,\, \mathbf{h}K_0\rangle_{\boldsymbol{H}(K_0)}\right]^2}{\langle \mathbf{h}K_0,\, \mathbf{h}K_0\rangle_{\boldsymbol{H}(K_0)}}$$

$$\leq \langle s_0,\, s_0\rangle_{\boldsymbol{H}(K_0)},$$

with equality if and only if $\mathbf{h}K_0 = cs_0$, $c \neq 0$. Further, since $s_0 \in \boldsymbol{H}(K_0)$, there exists a sequence $\{\bar{\mathbf{h}}_N^0\}_{N=1}^{\infty}$ of finite filters such that

$$\lim_{N\to\infty} \bar{\mathbf{h}}_N^0 K_0 = s_0,$$

and it follows that

$$\lim_{N\to\infty} \Delta(\bar{\mathbf{h}}_N^0;\,(s_0,K_0)) = \lim_{N\to\infty} \frac{\left[\langle s_0,\, \bar{\mathbf{h}}_N^0 K_0\rangle_{\boldsymbol{H}(K_0)}\right]^2}{\langle \bar{\mathbf{h}}_N^0 K_0,\, \bar{\mathbf{h}}_N^0 K_0\rangle_{\boldsymbol{H}(K_0)}}$$

$$= \langle s_0,\, s_0\rangle_{\boldsymbol{H}(K_0)}.$$

This proves the lemma. ■

The following theorem gives necessary and sufficient conditions for a pair $(\mathbf{h}_R;\,(s_L,K_L))$ to be a saddle point for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})),\boldsymbol{U},\Delta)$. It should be noted that, except for the addition of a scale factor on the robust filter $\mathbf{h}_R$, these conditions are identical to those given in Theorem 4.3.4 for the robust L^2-estimation problem.

Theorem 4.4.4: $(\mathbf{h}_R;\,(s_L,K_L)) \in \mathbb{F}(\mathbb{K}(\boldsymbol{U})) \times \boldsymbol{U}$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})),\boldsymbol{U},\Delta)$ if and only if $\mathbf{h}_R K_L = cs_L$ for some $c \neq 0$, and

$$2\langle s,\, \mathbf{h}'_R\rangle - \langle s_L,\, \mathbf{h}'_R\rangle - \langle \mathbf{h}'_R K,\, \mathbf{h}'_R\rangle \geq 0, \qquad \forall\ (s,K) \in \boldsymbol{U}, \tag{4.4.7}$$

where $\mathbf{h}'_R = \frac{1}{c}\mathbf{h}_R$.

Proof: It follows from (4.4.6) that $(\mathbf{h}_R; (s_L, K_L))$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\mathbb{U})), \mathbb{U}, \Delta)$ if and only if

$$\Delta(\mathbf{h}; (s_L, K_L)) \le \Delta(\mathbf{h}_R; (s_L, K_L)), \qquad \forall\ \mathbf{h} \in \mathbb{F}(\mathbb{K}(\mathbb{U})), \tag{4.4.8}$$

and

$$\Delta(\mathbf{h}_R; (s_L, K_L)) \le \Delta(\mathbf{h}_R; (s, K)), \qquad \forall\ (s, K) \in \mathbb{U}. \tag{4.4.9}$$

Lemma 4.4.3 implies that (4.4.8) is satisfied if and only if $\mathbf{h}_R K_L = c s_L$ for $c \neq 0$, so we need to establish that, given $\mathbf{h}_R K_L = c s_L$, (4.4.9) holds if and only if (4.4.7) holds. Since Δ is convex on $\mathbb{U}$ for fixed $\mathbf{h} \in \mathbb{F}(\mathbb{K}(\mathbb{U}))$, (4.4.7) holds if and only if, for every $(s, K) \in \mathbb{U}$,

$$\lim_{\alpha \to 0} \left[\Delta(\mathbf{h}_R; (s_\alpha, K_\alpha)) - \Delta(\mathbf{h}_R; (s_L, K_L))\right] \ge 0$$

where (s_α, K_α) is defined in the usual way. But if $\mathbf{h}_R K_L = c s_L$, then

$$\Delta(\mathbf{h}_R; (s_L, K_L)) = \langle s_L, s_L \rangle_{\mathbb{H}(K_L)} = \frac{1}{c} \langle s_L, \mathbf{h}_R \rangle = \frac{1}{c^2} \langle \mathbf{h}_R K_L, \mathbf{h}_R \rangle,$$

and

$$\lim_{\alpha\to 0}\frac{1}{\alpha}\Big[\Delta(\mathbf{h}_R;(s_\alpha,K_\alpha)) - \Delta(\mathbf{h}_R;(s_L,K_L))\Big] = \lim_{\alpha\to 0}\frac{1}{\alpha}\left[\frac{\left[\langle s_\alpha, \mathbf{h}_R\rangle\right]^2}{\langle \mathbf{h}_R K_\alpha, \mathbf{h}_R\rangle} - \frac{1}{c}\langle s_L, \mathbf{h}_R\rangle\right]$$

$$= \lim_{\alpha\to 0}\frac{1}{\alpha}\left[\frac{\alpha^2\left[\langle s, \mathbf{h}_R\rangle\right]^2 + 2\alpha(1-\alpha)\langle s, \mathbf{h}_R\rangle\langle s_L, \mathbf{h}_R\rangle + (1-\alpha)^2\left[\langle s_L, \mathbf{h}_R\rangle\right]^2}{\langle \mathbf{h}_R K_\alpha, \mathbf{h}_R\rangle} - \frac{1}{c}\langle s_L, \mathbf{h}_R\rangle\right]$$

$$= \frac{2}{c}\langle s, \mathbf{h}_R\rangle - \frac{2}{c}\langle s_L, \mathbf{h}_R\rangle + \lim_{\alpha\to 0}\frac{1}{c\alpha}\left[\langle s_L, \mathbf{h}_R\rangle - \frac{1}{c}\langle \mathbf{h}_R K_\alpha, \mathbf{h}_R\rangle\right]$$

$$= \frac{2}{c}\langle s, \mathbf{h}_R\rangle - \frac{2}{c}\langle s_L, \mathbf{h}_R\rangle + \lim_{\alpha\to 0}\frac{1}{c^2\alpha}\left[\alpha\langle \mathbf{h}_R K_L, \mathbf{h}_R\rangle - \alpha\langle \mathbf{h}_R K, \mathbf{h}_R\rangle\right]$$

$$= \frac{2}{c}\langle s, \mathbf{h}_R\rangle - \frac{1}{c}\langle s_L, \mathbf{h}_R\rangle - \frac{1}{c^2}\langle \mathbf{h}_R K, \mathbf{h}_R\rangle$$

$$= 2\langle s, \mathbf{h}'_R\rangle - \langle s_L, \mathbf{h}'_R\rangle - \langle \mathbf{h}'_R K, \mathbf{h}'_R\rangle.$$

Hence, if $\mathbf{h}_R K_L = c s_L$, (4.4.9) holds if and only if (4.4.7) holds, and the theorem follows. ■

We also have the following corollary, which is analogous to Corollary 4.3.6.

Corollary 4.4.5: Suppose that (s_L, K_L) is a least favorable pair for $(\mathbb{F}(\mathbb{K}(\mathbb{U})), \mathbb{U}, \Delta)$ and K_L dominates $\mathbb{K}(\mathbb{U})$. Let $\mathbf{h}_R$ be a matched filter for (s_L, K_L); that is, $\mathbf{h}_R K_L = c s_L$ for some $c \neq 0$. Then $(\mathbf{h}_R; (s_L, K_L))$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\mathbb{U})), \mathbb{U}, \Delta)$.

Proof: Let $\mathbf{h}'_R = \frac{1}{c}\mathbf{h}_R$ so that $\mathbf{h}'_R K_L = s_L$. By hypothesis, (s_L, K_L) is a least favorable pair, so, for all $(s, K) \in \mathbb{U}$,

$$\Delta^*(s_L, K_L) = \langle s_L, s_L\rangle_{\mathbb{H}(K_L)} \le \langle s, s\rangle_{\mathbb{H}(K)} = \Delta^*(s, K).$$

Since K_L dominates $\mathbb{K}(\mathbb{U})$, it follows from Theorem 4.3.5 that $\mathbf{h}'_R \in \mathbb{F}(\mathbb{K}(\mathbb{U}))$ and

$$2\langle s, \mathbf{h}'_R\rangle - \langle s_L, \mathbf{h}'_R\rangle - \langle \mathbf{h}'_R K, \mathbf{h}'_R\rangle \geq 0, \qquad \forall\ (s, K) \in \boldsymbol{U}.$$

The result follows by applying Theorem 4.4.4. ■

At this point, it is clear that the characteristics of a saddle point solution to the robust matched filtering problem are virtually identical to those of a saddle point solution to the robust L^2-estimation problem. Indeed, if we insist upon using normalized matched filters (i.e., $\mathbf{h}_R K_L = s_L$), the characteristics are the same. This being the case, we can merely observe that, with the exception of Corollary 4.3.9 (the proof of which depends explicitly upon the structure of the function M), all of the remarks, corollaries, and examples following Corollary 4.3.6 have exact analogs relating to the robust matched filtering problem. In particular, if the covariance structure of the observed process is assumed to be known, so that the only uncertainty is in the structure of the signal m, the existence of a robust matched filter is guaranteed.

With reference to the case in which the covariance function K_X is assumed to be known, it is interesting to note that the matched filter corresponding to any signal $m \in \boldsymbol{H}(K_X)$ will be a robust matched filter for any admissible uncertainty class $\boldsymbol{U}$ in which the least favorable signal is a multiple of m. This follows immediately from Corollary 4.4.5 and holds, in particular, for the class $\boldsymbol{U}$ given by

$$\boldsymbol{U} \triangleq \left\{(s, K_X) \colon \|s - m\|^2_{\boldsymbol{H}(K_X)} \leq \varepsilon\right\},$$

for any $0 < \varepsilon < \|m\|^2_{\boldsymbol{H}(K_X)}$. This is a generalization of a more familiar result regarding matched filters for signals in white noise (see [25], §III.A).[3]

[3] A similar result has been noted recently by Donoho and Liu in a different robustness framework [13].

Remark 4.4.6: For the robust matched filtering problem, the requirement that $s \in \boldsymbol{H}(K)$ for all $(s, K) \in \boldsymbol{U}$ is not as natural as it is for the L^2-estimation problem. Although such an uncertainty class does usually admit a rich class of signal-covariance pairs, it is sometimes the case that the "natural" uncertainty class contains pairs for which $s \notin \boldsymbol{H}(K)$. For example, if I is a Hilbert space $\boldsymbol{H}_0$ and K_0 is a known covariance operator on $\boldsymbol{H}_0$, a common uncertainty class corresponding to signal uncertainty is of the form $S \times \{K_0\}$, where

$$S \triangleq \left\{ s \in \boldsymbol{H}_0 : \|s - s_0\|^2_{\boldsymbol{H}_0} \leq \varepsilon \right\}$$

for some nominal signal s_0 and $0 < \varepsilon < \|s_0\|^2_{\boldsymbol{H}_0}$ (see [53]). If K_0 is not an invertible operator, such a set will contain signals that are not in $\boldsymbol{H}(K_0)$.

Consider the problem of finding robust matched filters for the case in which the uncertainty class $\boldsymbol{U}$ contains pairs (s, K) for which $s \notin \boldsymbol{H}(K)$. Naturally, we must restrict our attention to the set of filters $\boldsymbol{F}(\boldsymbol{U}) \subset \boldsymbol{F}(\mathbb{K}(\boldsymbol{U}))$ that are well-defined for all signals "contained in" $\boldsymbol{U}$. That is, $\mathbf{h} = \{\tilde{\mathbf{h}}_N\}_{N=1}^{\infty}$ is a member of $\boldsymbol{F}(\boldsymbol{U})$ if, for every $(s, K) \in \boldsymbol{U}$, $\mathbf{h} \in \boldsymbol{F}(K)$, and

$$\langle s, \mathbf{h} \rangle \triangleq \lim_{N \to \infty} \langle s, \tilde{\mathbf{h}}_N \rangle \triangleq \lim_{N \to \infty} \sum_{i=1}^{\infty} h_i(N) s(t_i)$$

is well-defined. If we seek saddle points for this problem, then, as the following lemma shows, under mild regularity conditions, we can identify all possible saddle points by restricting attention to those pairs $(s, K) \in \boldsymbol{U}$ for which $s \in \boldsymbol{H}(K)$. The proof of Lemma 4.4.7 is a straightforward application of the RKHS Approximation Lemma, which is given in Appendix C.

Lemma 4.4.7: Let I be a separable metric space and $\boldsymbol{U}$ a convex set of pairs (s, K) such that K is a continuous covariance function on I and $s : I \to \mathbb{R}$ is continuous. Then

$$\sup_{\mathbf{h} \in \boldsymbol{F}(\boldsymbol{U})} \Delta(\mathbf{h}; (s, K)) \triangleq \sup_{\mathbf{h} \in \boldsymbol{F}(\boldsymbol{U})} \frac{\left\{\langle s, \mathbf{h} \rangle\right\}^2}{\langle \mathbf{h}K, \mathbf{h} \rangle} < \infty$$

if and only if $s \in \boldsymbol{H}(K)$. Defining the associated admissible uncertainty class $\boldsymbol{U}'$ by

$$\boldsymbol{U}' \triangleq \left\{(s, K) \in \boldsymbol{U} : s \in \boldsymbol{H}(K)\right\}$$

and assuming that $\boldsymbol{U}'$ is nonempty, it follows that any least favorable pair (s_L, K_L) for $(\boldsymbol{F}(\boldsymbol{U}), \boldsymbol{U}, \Delta)$ must satisfy $(s_L, K_L) \in \boldsymbol{U}'$. Hence, $(\mathbf{h}_R; (s_L, K_L))$ is a saddle point for $(\boldsymbol{F}(\boldsymbol{U}), \boldsymbol{U}, \Delta)$ only if it is also a saddle point for $(\boldsymbol{F}(\mathbb{K}(\boldsymbol{U}')), \boldsymbol{U}', \Delta)$.

Note that the regularity conditions for Lemma 4.4.7 are not very restrictive. They are satisfied, for example, if I is a countable set (endowed with the trivial metric), or if I is a separable Hilbert space $\boldsymbol{H}_0$, and $\boldsymbol{U}$ corresponds to a collection of pairs (s, K) with $s \in \boldsymbol{H}_0$ and K a bounded covariance operator on $\boldsymbol{H}_0$. Also note that Lemma 4.4.7 does *not* say that if $(\mathbf{h}_R; (s_L, K_L)) \in \boldsymbol{F}(\boldsymbol{U}) \times \boldsymbol{U}'$ is a saddle point for $(\boldsymbol{F}(\mathbb{K}(\boldsymbol{U}')), \boldsymbol{U}', \Delta)$, then it is necessarily a saddle point for $(\boldsymbol{F}(\boldsymbol{U}), \boldsymbol{U}, \Delta)$. However, in many problems of interest, this will in fact be true, as the following lemma indicates. The proof of Lemma 4.4.8 is given in Appendix C.

Lemma 4.4.8 Let I be a Hilbert space $\boldsymbol{H}_0$ and let $\boldsymbol{U} = S \times \mathbb{K}$, where $S \subset \boldsymbol{H}_0$ is convex with a nonempty interior and $\mathbb{K}$ is a convex class of covariance operators on $\boldsymbol{H}_0$. Let $\boldsymbol{U}' \triangleq \{(s, \mathrm{K}) \in \boldsymbol{U} : s \in \boldsymbol{H}(\mathrm{K})\}$. If there exists $h_R \in \boldsymbol{H}_0$ and $(s_L, \mathrm{K}_L) \in \boldsymbol{U}'$ such that K_L is strictly positive and $(h_R; (s_L, K_L))$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\mathbb{U}')), \mathbb{U}', \Delta)$, then $(h_R; (s_L, K_L))$ is a saddle point for $(\mathbb{F}(\mathbb{U}), \mathbb{U}, \Delta)$.

We close this section with a simple example.

Example 4.4.9: As in Example 4.3.10, let $I = L^2([a,b])$, k_0 be a continuous covariance function on $[a,b]$ with associated strictly positive covariance operator K_0, and $s_0 \in \mathbb{H}(K_0)$. Let $\mathbb{U} = S \times \{K_0\}$, where

$$S \triangleq \left\{ s \in L^2([a,b]) : \int_a^b [s(\tau) - s_0(\tau)]^2 d\tau \le \varepsilon \right\}$$

for some $0 < \varepsilon < \int_a^b s_0^2(\tau) d\tau$. As discussed above, we define $\mathbb{U}'$ as

$$\mathbb{U}' \triangleq \left\{ (s, K_0) \in \mathbb{U} : s \in \mathbb{H}(K_0) \right\},$$

and it is obvious that $\mathbb{U}' = S' \times \{K_0\}$, where

$$S' \triangleq \left\{ s \in \mathbb{H}(K_0) : \int_a^b [s(\tau) - s_0(\tau)]^2 d\tau \le \varepsilon \right\}.$$

Proceeding exactly as in Example 4.3.10, we find that $(h_R; (s_L, K_0))$ is a saddle point for $(\mathbb{F}(K_0), \mathbb{U}', \Delta)$, where s_L and h_R are defined by (4.3.14) and (4.3.17), respectively. It follows from Lemma 4.4.8 that $(h_R; (s_L, K_0))$ is also a saddle point for $(\mathbb{F}(\mathbb{U}), \mathbb{U}, \Delta)$.[4]

[4]This generalizes an example given in [53], in which it was shown that h_R is a robust matched filter for the game $(L^2([a,b]), \mathbb{U}, \Delta)$. In our terminology, this corresponds to restricting the class of allowable filters to include only finite filters.

4.5. Robust Quadratic Detection

Throughout this section, we assume that the observed process $\mathrm{X} \triangleq \{X(t);\ t \in I\}$ is Gaussian with mean zero. Consider the hypothesis testing problem:

$$\begin{aligned} &\mathrm{H}_0\text{: X has covariance function } K_{\mathrm{N}} \\ \text{versus} \qquad\qquad &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (4.5.1) \\ &\mathrm{H}_1\text{: X has covariance function } K_{\mathrm{N}} + K_{\mathrm{S}}, \end{aligned}$$

where K_{S} and K_{N} are known covariance functions on I and $K_{\mathrm{S}} \in \boldsymbol{H}(K_{\mathrm{N}}) \otimes \boldsymbol{H}(K_{\mathrm{N}})$. This corresponds to the problem of detecting a zero-mean Gaussian signal with covariance function K_{S} in the presence of additive, independent, zero-mean Gaussian noise with covariance function K_{N}. Here again, the condition $K_{\mathrm{S}} \in \boldsymbol{H}(K_{\mathrm{N}}) \otimes \boldsymbol{H}(K_{\mathrm{N}})$ is a regularity condition and is related to the nonsingularity of (4.5.1), as discussed in Chapter 2. In particular, this condition guarantees that there exists a constant $0 < C < \infty$ such that $K_{\mathrm{S}} \ll CK_{\mathrm{N}}$; that is, K_{N} dominates K_{S} (see [1], §I.11).

We assume that Problem (4.5.1) is to be decided using a quadratic detector; that is, a detector in which the test statistic $\Phi(\mathrm{X})$ is of the form

$$\Phi(\mathrm{X}) = \lim_{N \to \infty} \tilde{\Phi}_N(\mathrm{X}) \triangleq \lim_{N \to \infty} \sum_{i,j=1}^{\infty} \phi_{ij}(N) X(t_i) X(t_j), \tag{4.5.2}$$

where, for each N, the infinite-dimensional real matrix $(\phi_{ij}(N))$ is symmetric and has only finitely many nonzero elements. The limit is taken in the mean-square sense under both hypotheses; that is,

$$\lim_{N \to \infty} \mathrm{E}_n \{[\Phi(\mathrm{X}) - \bar{\Phi}_N(\mathrm{X})]^2\} = 0, \qquad n = 0 \text{ and } 1. \tag{4.5.3}$$

Note that this definition is very general and includes the more conventional examples of quadratic statistics as special cases. Since K_{N} dominates K_{S}, the limits in (4.5.3) exist if and only if (see [3] for details)

$$\begin{aligned} 0 &= \lim_{N,M \to \infty} \mathrm{Var}_0\{\bar{\Phi}_N(\mathrm{X}) - \bar{\Phi}_M(\mathrm{X})\} \\ &= \lim_{N,M \to \infty} \mathrm{Var}_0\left\{\sum_{i,j=1}^{\infty} [\phi_{ij}(N) - \phi_{ij}(M)] X(t_i) X(t_j)\right\} \\ &= 2\left\{\lim_{N,M \to \infty} \sum_{i,j=1}^{\infty} \sum_{k,l=1}^{\infty} [\phi_{ij}(N) - \phi_{ij}(M)][\phi_{kl}(N) - \phi_{kl}(M)] K_{\mathrm{N}}(t_k,t_i) K_{\mathrm{N}}(t_l,t_j)\right\}, \end{aligned} \tag{4.5.4}$$

and

$$\begin{aligned} 0 &= \lim_{N,M \to \infty} \mathrm{E}_0\{\bar{\Phi}_N(\mathrm{X}) - \bar{\Phi}_M(\mathrm{X})\} \\ &= \lim_{N,M \to \infty} \left\{\sum_{i,j=1}^{\infty} [\phi_{ij}(N) - \phi_{ij}(M)] K_{\mathrm{N}}(t_j,t_i)\right\}. \end{aligned} \tag{4.5.5}$$

Recall that the space $\boldsymbol{H}(K_{\mathrm{N}}) \otimes \boldsymbol{H}(K_{\mathrm{N}})$ is equivalent to the RKHS $\boldsymbol{H}(K_{\mathrm{N}}^2)$, where the reproducing kernel $K_{\mathrm{N}}^2 : I^2 \times I^2 \to \mathbb{R}$ is given by

$$K_{\mathrm{N}}^2((t_1,\tau_1);\,(t_2,\tau_2)) \triangleq K_{\mathrm{N}}(t_1,t_2) K_{\mathrm{N}}(\tau_1,\tau_2), \qquad (t_i,\tau_i) \in I^2.$$

If we define $\bar{\Phi}_N K_{\mathrm{N}}^2 \in \boldsymbol{H}(K_{\mathrm{N}}^2)$ by

$$\begin{aligned} \bar{\Phi}_N K_{\mathrm{N}}^2(\cdot,*) &\triangleq \sum_{i,j=1}^{\infty} \phi_{ij}(N) K_{\mathrm{N}}^2((\cdot,*);\,(t_i,t_j)) \\ &= \sum_{i,j=1}^{\infty} \phi_{ij}(N) K_{\mathrm{N}}(\cdot,t_i) K_{\mathrm{N}}(*,t_j), \end{aligned}$$

then it follows from (4.5.4) that

$$\lim_{N,M \to \infty} \|\tilde{\Phi}_N K_{\mathrm{N}}^2 - \tilde{\Phi}_M K_{\mathrm{N}}^2\|^2_{\boldsymbol{H}(K_{\mathrm{N}}^2)} = 0,$$

and we can define $\Phi K_{\mathrm{N}}^2 \in \boldsymbol{H}(K_{\mathrm{N}}^2)$ as

$$\Phi K_{\mathrm{N}}^2(\cdot,*) \triangleq \lim_{N \to \infty} \tilde{\Phi}_N K_{\mathrm{N}}^2(\cdot,*) = \lim_{N \to \infty} \sum_{i,j=1}^{\infty} \phi_{ij}(N) K_{\mathrm{N}}(\cdot,t_i) K_{\mathrm{N}}(*,t_j).$$

Clearly then, corresponding to any statistic $\Phi(\mathrm{X})$ of the form (4.5.2), there exists a *symmetric* filter $\Phi \in \mathbb{F}(K_{\mathrm{N}}^2)$. The term symmetric refers to the fact that $\Phi = \{\tilde{\Phi}_N\}_{N=1}^{\infty}$, where, for each finite filter $\tilde{\Phi}_N$, the matrix $(\phi_{ij}(N))$ is symmetric. Conversely, given an arbitrary symmetric filter $\Phi \in \mathbb{F}(K_{\mathrm{N}}^2)$, there exists a corresponding statistic $\Phi(\mathrm{X})$ of the form (4.5.2) only if (4.5.5) is satisfied. This need not always be the case, but for the remainder of this section, we will ignore any possible convergence problems and simply assume that (4.5.5) is satisfied for any filter under consideration. In any case, for the performance criterion considered here (or, for that matter, most other common criteria), the satisfaction of (4.5.5) is unrelated to the performance characteristics of the filter.

Quadratic detectors for deciding Problem (4.5.1) are often compared using the *deflection ratio*, which is one member of the class of so-called *generalized signal-to-noise ratios* (see [17] for a discussion of this class of performance measures). The deflection ratio for Problem (4.5.1) corresponding to a test statistic $\Phi(\mathrm{X})$ of the form (4.5.2) with associated symmetric filter $\Phi \in \mathbb{F}(K_{\mathrm{N}}^2)$ is given by

$$D(\Phi;\,(K_S,K_N^2)) \triangleq \frac{\left[E_1\{\Phi(X)\} - E_0\{\Phi(X)\}\right]^2}{\mathrm{Var}_0\{\Phi(X)\}}$$

$$= \frac{\left[\lim_{N\to\infty} \sum_{i,j=1}^{\infty} \phi_{ij}(N) K_S(t_j,t_i)\right]^2}{2\left[\lim_{N\to\infty} \sum_{i,j=1}^{\infty} \sum_{k,l=1}^{\infty} \phi_{ij}\phi_{kl} K_N(t_k,t_i) K_N(t_l,t_j)\right]}$$

$$= \frac{\left[\left\langle K_S,\, \Phi K_N^2 \right\rangle_{\mathbb{H}(K_N^2)}\right]^2}{2\left\langle \Phi K_N^2,\, \Phi K_N^2 \right\rangle_{\mathbb{H}(K_N^2)}}.$$

We extend the definition of D to include arbitrary filters in the obvious way; that is, for any $\Phi \in \mathbb{F}(K_N^2)$, we define

$$D(\Phi;\,(K_S,K_N^2)) \triangleq \frac{\left[\left\langle K_S,\, \Phi K_N^2 \right\rangle_{\mathbb{H}(K_N^2)}\right]^2}{2\left\langle \Phi K_N^2,\, \Phi K_N^2 \right\rangle_{\mathbb{H}(K_N^2)}}. \tag{4.5.6}$$

Comparing with (4.4.2), we see that the deflection ratio takes the form of a signal-to-noise ratio, where we regard I^2 as the index set, $K_S\colon I^2 \to \mathbb{R}$ as the signal, and $K_N^2 : I^2 \times I^2 \to \mathbb{R}$ as the covariance function.

It follows immediately from (4.5.6) and the Schwarz inequality that a filter $\Phi^* \in \mathbb{F}(K_N^2)$ satisfies

$$D(\Phi^*;\,(K_S,K_N^2)) = \sup_{\Phi\in\mathbb{F}(K_N^2)} D(\Phi;\,(K_S,K_N^2)) \tag{4.5.7}$$

if and only if $\Phi^* K_N^2 = cK_S$ for some $c \neq 0$, and the maximum possible deflection ratio for Problem (4.5.1) is given by

$$D(\Phi^*;\,(K_S,K_N^2)) = \tfrac{1}{2}\|K_S\|^2_{\mathbb{H}(K_N^2)}.$$

Since $K_S \in \boldsymbol{H}(K_N) \otimes \boldsymbol{H}(K_N) = \boldsymbol{H}(K_N^2)$, a filter $\Phi^* \in \boldsymbol{F}(K_N^2)$ satisfying (4.5.7) can always be found. Further, if we let Φ^T represent the "transpose" of a filter $\Phi \in \boldsymbol{F}(K_N^2)$, defined in the obvious way, then $\Phi^T \in \boldsymbol{F}(K_N^2)$, and $\Phi K_N^2 = cK_S$ if and only if $\Phi^T K_N^2 = cK_S$. Hence, a filter $\Phi^* \in \boldsymbol{F}(K_N^2)$ satisfies (4.5.7) only if $\Phi^* + (\Phi^*)^T$ does, and it follows that a symmetric filter satisfying (4.5.7) always exists.

Now suppose that the pair (K_S, K_N) is known only to belong to some set and let $\boldsymbol{U}$ represent the associated set of "signal-covariance" pairs (s, K), $s \in \boldsymbol{H}(K)$, corresponding to the possible values of (K_S, K_N^2). The set $\boldsymbol{K}(\boldsymbol{U})$ is then the set of covariance functions $K : I^2 \times I^2 \rightarrow \boldsymbol{R}$ corresponding to the possible values of K_N^2. To decide the stochastic signal detection problem in the presence of these uncertainties, one might wish to use a quadratic detector incorporating a symmetric filter $\Phi_R \in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}))$ that satisfies

$$\inf_{(s,K)\in \boldsymbol{U}} D(\Phi_R; (s,K)) = \sup_{\Phi\in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}))} \inf_{(s,K)\in \boldsymbol{U}} D(\Phi; (s,K)). \tag{4.5.8}$$

We refer to an arbitrary filter $\Phi_R \in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}))$ satisfying (4.5.8) as a *robust filter* for the game $(\boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U})), \boldsymbol{U}, D)$. If Φ_R is also symmetric, we refer to it as a *robust symmetric filter*.

As usual, to find a robust filter we search for a *saddle point* $(\Phi_R; (s_L, K_L)) \in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U})) \times \boldsymbol{U}$ satisfying

$$D(\Phi; (s_L, K_L)) \leq D(\Phi_R; (s_L, K_L)) \leq D(\Phi_R; (s,K)), \qquad \forall\ \Phi \in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U})),\ (s,K) \in \boldsymbol{U}. \tag{4.5.9}$$

It follows immediately (as in Lemma 4.4.3) that $(\Phi_R; (s_L, K_L))$ is a saddle point for $(\boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U})), \boldsymbol{U}, D)$ only if $\Phi_R K_L = cs_L$, $c \neq 0$, and

$$D^*(s_L, K_L) \leq D^*(s,K), \qquad \forall\ (s,K) \in \boldsymbol{U},$$

where

$$D^*(s,K) \triangleq \sup_{\Phi \in \boldsymbol{F}(\mathbb{K}(\boldsymbol{U}))} D(\Phi; (s,K)) = \tfrac{1}{2}\|s\|^2_{\boldsymbol{H}(K)}. \tag{4.5.10}$$

Further, it is easy to see that, for any $\Phi \in \boldsymbol{F}(\mathbb{K}(\boldsymbol{U}))$,

$$D(\Phi; (s,K)) \le D((\Phi + \Phi^T); (s,K)), \qquad \forall\, (s,K) \in \boldsymbol{U},$$

which implies that $(\Phi_R; (s_L,K_L))$ is a saddle point for $(\boldsymbol{F}(\mathbb{K}(\boldsymbol{U})),\boldsymbol{U},D)$ only if $((\Phi_R + \Phi_R^T); (s_L,K_L))$ is also a saddle point. Hence, robust filters corresponding to saddle points for $(\boldsymbol{F}(\mathbb{K}(\boldsymbol{U})),\boldsymbol{U},D)$ can always be taken to be symmetric.

Now suppose that the noise covariance K_N is assumed to be known but that K_S is known only to belong to a convex set $S \subset \boldsymbol{H}(K_N^2)$. The set

$$\boldsymbol{U} \triangleq \left\{(s,K_N^2): s \in S\right\}$$

is then an admissible uncertainty class defined on the index set I^2. Hence, the problem of finding robust filters for $(\boldsymbol{F}(K_N^2),\boldsymbol{U},D)$ is identical in form to that of finding robust matched filters when the noise covariance is known and the deterministic signal belongs to some convex set. It follows by an argument analogous to that given in Corollary 4.3.8 that a robust symmetric filter for $(\boldsymbol{F}(K_N^2),\boldsymbol{U},D)$ exists and can be chosen to satisfy $\Phi_R K_N^2 = s_L$, where s_L is the unique element of $\overline{S}$ (the closure of S in $\boldsymbol{H}(K_N^2)$) with minimum norm.

If K_N is not assumed to be known, the situation is somewhat more complicated. In this case, the set $\boldsymbol{U}$ will not generally be convex, and we cannot apply the results of previous sections. However, if we let $\boldsymbol{U}' \triangleq \mathrm{co}(\boldsymbol{U})$ (the convex hull of $\boldsymbol{U}$), then it is easy to see that $\boldsymbol{U}'$ is an admissible uncertainty class on the index set I^2. The problem of finding robust filters for the game $(\mathbb{F}(\mathbb{K}(\boldsymbol{U}')),\boldsymbol{U}',D)$ thus takes the form of the robust matched filtering problem discussed in Section 4.4, and we can apply previous results to the search for saddle points. Since $\boldsymbol{U}' = \mathrm{co}(\boldsymbol{U})$, we again have, for all $\Phi \in \mathbb{F}(\mathbb{K}(\boldsymbol{U}'))$,

$$D(\Phi; (s,K)) \le D((\Phi + \Phi^T); (s,K)), \qquad \forall\, (s,K) \in \boldsymbol{U}',$$

and it follows that robust filters corresponding to saddle points for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U}')),\boldsymbol{U}',D)$ can be taken to be symmetric.

It is true, of course, that a saddle point for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U}')),\boldsymbol{U}',D)$ need not be a saddle point for the original problem $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})),\boldsymbol{U},D)$. However, if $(\Phi_R; (s_L,K_L))$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U}')),\boldsymbol{U}',D)$, then it follows from (4.5.9) and (4.5.10) that

$$0 < \tfrac{1}{2}\|s_L\|^2_{\boldsymbol{H}(K_L)} = D(\Phi_R; (s_L,K_L)) \le D(\Phi_R; (s,K)), \qquad \forall\, (s,K) \in \boldsymbol{U}.$$

Hence, the performance of the filter Φ_R (as measured by the deflection ratio) is bounded below by $\tfrac{1}{2}\|s_L\|^2_{\boldsymbol{H}(K_L)}$ for all $(s,K) \in \boldsymbol{U}$. In this respect, Φ_R remains robust for the smaller problem $(\mathbb{F}(\mathbb{K}(\boldsymbol{U})),\boldsymbol{U},D)$ although it is possible that there are yet other filters whose worst-case performance over $\boldsymbol{U}$ strictly exceeds that of Φ_R.

To illustrate the robust quadratic detection problem, we consider the following simple example in which the covariance function of the stochastic signal is assumed to be known. For the sake of convenience, we consider a set $\boldsymbol{U}'$ that is larger than $\mathrm{co}(\boldsymbol{U})$.

Example 4.5.1: Let $I = \mathbb{Z}$, the set of integers and let $\Omega = [-\pi,\pi]$. Suppose that the zero-mean, Gaussian signal has known covariance function K_S satisfying

$$\sum_{i,j=-\infty}^{\infty} [K_S(i,j)]^2 < \infty.$$

Note that this implies that the signal is a nonstationary process. Let σ be the (necessarily symmetric) two-dimensional Fourier transform of the covariance function K_S; that is, $\sigma \in L^2(\Omega^2)$ and

$$K_S(m,n) = \frac{1}{4\pi^2} \int_\Omega \int_\Omega e^{i\omega m} e^{i\lambda n} \sigma(\omega,\lambda) d\omega d\lambda, \qquad \forall\ (m,n) \in \mathbb{Z}^2.$$

Suppose that the noise process is a zero-mean stationary Gaussian process with covariance function K_N having power spectral density ν that is known only to satisfy

$$\frac{1}{2\pi} \int_\Omega \nu(\omega) d\omega = \rho, \tag{4.5.11}$$

and

$$\nu_l(\omega) \le \nu(\omega) \le \nu_u(\omega), \qquad \forall\ \omega \in \Omega, \tag{4.5.12}$$

where ν_l and ν_u are known functions satisfying

$$0 < \inf_{\omega \in \Omega} \nu_l(\omega) < \sup_{\omega \in \Omega} \nu_u(\omega) < \infty,$$

and

$$\frac{1}{2\pi} \int_\Omega \nu_l(\omega) d\omega < \rho < \frac{1}{2\pi} \int_\Omega \nu_u(\omega) d\omega.$$

Corresponding to any power spectral density ν satisfying (4.5.11) and (4.5.12), we have

$$K_N^2((m_1,n_1);(m_2,n_2)) = K_N(m_1,m_2) K_N(n_1,n_2) = \frac{1}{4\pi^2} \int_\Omega \int_\Omega e^{i\omega(m_1-m_2)} e^{i\lambda(n_1-n_2)} \nu(\omega)\nu(\lambda) d\omega d\lambda,$$

and it follows that the set $\mathcal{U}$ is given by

$$\boldsymbol{U} \triangleq \Bigg\{(s,K):\ s = K_S \text{ and } K((m_1,n_1);\ (m_2,n_2)) = \frac{1}{4\pi^2}\iint_{\Omega\Omega} e^{i\omega(m_1-m_2)}e^{i\lambda(n_1-n_2)}\nu(\omega)\nu(\lambda)\,d\omega\, d\lambda,$$

$$\text{where } \nu_l(\omega) \le \nu(\omega) \le \nu_u(\omega) \text{ and } \frac{1}{2\pi}\int_{\Omega}\nu(\omega)d\omega = \rho\Bigg\}.$$

For this problem then, it is convenient to take the set $\boldsymbol{U}' \supset \text{co}(\boldsymbol{U})$ to be

$$\boldsymbol{U}' \triangleq \Bigg\{(s,K):\ s = K_S \text{ and } K((m_1,n_1);\ (m_2,n_2)) = \frac{1}{4\pi^2}\iint_{\Omega\Omega} e^{i\omega(m_1-m_2)}e^{i\lambda(n_1-n_2)}\eta(\omega,\lambda)\,d\omega\, d\lambda,$$

$$\text{where } \nu_l(\omega)\nu_l(\lambda) \le \eta(\omega,\lambda) \le \nu_u(\omega)\nu_u(\lambda) \text{ and } \frac{1}{4\pi^2}\iint_{\Omega\Omega}\eta(\omega,\lambda)d\omega = \rho^2\Bigg\},$$

and we seek a robust symmetric filter Φ_R for the game $(\boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}')),\boldsymbol{U}',D)$.

Note that every $K \in \boldsymbol{K}(\boldsymbol{U}')$ takes the form of a covariance function for a two-dimensional stationary random field with power spectral density η satisfying

$$\nu_l(\omega)\nu_l(\lambda) \le \eta(\omega,\lambda) \le \nu_u(\omega)\nu_u(\lambda), \qquad \forall\ (\omega,\lambda) \in \Omega^2, \tag{4.5.13}$$

and

$$\frac{1}{4\pi^2}\iint_{\Omega\Omega}\eta(\omega,\lambda)d\omega d\lambda = \rho^2. \tag{4.5.14}$$

Further, for every $K \in \boldsymbol{K}(\boldsymbol{U}')$ with power spectral density η, the RKHS $\boldsymbol{H}(K)$ consists of functions of the form (see [48], §8)

$$f(m,n) = \frac{1}{4\pi^2}\iint_{\Omega\Omega} e^{i\omega m}e^{i\lambda n}F(\omega,\lambda)\eta(\omega,\lambda)d\omega d\lambda, \qquad \forall\ (m,n) \in \mathbb{Z}^2, \tag{4.5.15}$$

where

$$\|f\|^2_{\boldsymbol{H}(K)} \triangleq \frac{1}{4\pi^2}\iint_{\Omega\Omega}|F(\omega,\lambda)|^2\eta(\omega,\lambda)d\omega d\lambda < \infty. \tag{4.5.16}$$

Also note that each $K \in \mathbb{K}(\boldsymbol{U}')$ dominates $\mathbb{K}(\boldsymbol{U}')$. In particular, if $K_0 \in \mathbb{K}(\boldsymbol{U}')$, then for any other $K \in \mathbb{K}(\boldsymbol{U}')$, we have

$$K \ll \left[\frac{\sup_{\omega \in \Omega} \nu_u(\omega)}{\inf_{\omega \in \Omega} \nu_l(\omega)} \right]^2 K_0.$$

It follows from Corollary 4.4.5 that $(\Phi_R; (s_L, K_L))$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U}')), \boldsymbol{U}', D)$ if and only if (s_L, K_L) is least favorable for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U}')), \boldsymbol{U}', D)$ and $\Phi_R K_L = c s_L$ for some $c \neq 0$. Clearly, in this case we will have $s_L = K_S$. Hence, if we can find $K_L \in \mathbb{K}(\boldsymbol{U}')$ satisfying

$$D^*(s_L, K_L) = \tfrac{1}{2} \|K_S\|^2_{\boldsymbol{H}(K_L)} \leq \tfrac{1}{2} \|K_S\|^2_{\boldsymbol{H}(K)} = D^*(s_L, K), \qquad \forall\, K \in \mathbb{K}(\boldsymbol{U}'), \tag{4.5.17}$$

then the filter Φ_R satisfying $\Phi_R K_L = K_S$ will be a robust filter for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U}')), \boldsymbol{U}', D)$.

To find K_L satisfying (4.5.17), we note that, for any $K \in \mathbb{K}(\boldsymbol{U}')$ with power spectral density η,

$$K_S(m,n) = \frac{1}{4\pi^2} \int_\Omega \int_\Omega e^{i\omega m} e^{i\lambda n} \frac{\sigma(\omega,\lambda)}{\eta(\omega,\lambda)} \eta(\omega,\lambda)\, d\omega\, d\lambda, \qquad \forall\, (m,n) \in \mathbb{Z}^2.$$

Since $\inf_{(\omega,\lambda) \in \Omega^2} \eta(\omega,\lambda) > 0$, it follows from (4.5.15) and (4.5.16) that $K_S \in \boldsymbol{H}(K)$ and

$$\|K_S\|^2_{\boldsymbol{H}(K)} = \frac{1}{4\pi^2} \int_\Omega \int_\Omega \frac{|\sigma(\omega,\lambda)|^2}{\eta(\omega,\lambda)}\, d\omega\, d\lambda. \tag{4.5.18}$$

Hence, finding $K_L \in \mathbb{K}(\boldsymbol{U}')$ satisfying (4.5.17) is equivalent to finding the function η_L minimizing (4.5.18) subject to the constraints (4.5.13) and (4.5.14). This problem has been solved in [24], where it is shown that η_L satisfies

$$\eta_L(\omega,\lambda) = \max\,\{\nu_l(\omega)\nu_l(\lambda),\ \min\,\{c\,|\sigma(\omega,\lambda)|,\ \nu_u(\omega)\nu_u(\lambda)\}\}, \qquad \forall\ (\omega,\lambda) \in \Omega^2,$$

if $c > 0$ can be found such that

$$\frac{1}{4\pi^2}\iint_{\Omega\Omega} \eta_L(\omega,\lambda)\,d\omega\, d\lambda = \rho^2.$$

Assuming that such a c exists,[5] we define the filter Φ_R as the infinite dimensional matrix with entries ϕ^R_{mn} given by

$$\phi^R_{mn} \triangleq \frac{1}{4\pi^2}\iint_{\Omega\Omega} e^{i\omega m} e^{i\lambda n} \frac{\sigma(\omega,\lambda)}{\eta_L(\omega,\lambda)}\, d\omega\, d\lambda, \qquad \forall\ (m,n) \in \mathbb{Z}^2.$$

Since $\inf\limits_{(\omega,\lambda)\in\Omega^2} \eta_L(\omega,\lambda) > 0$, $\dfrac{\sigma}{\eta_L} \in L^2(\Omega^2)$, and Φ_R is well-defined. Further, since $\dfrac{\sigma}{\eta_L}$ is symmetric, so is Φ_R. Finally, for all $(m,n) \in \mathbb{Z}^2$, we have

$$\begin{aligned}
\Phi_R K_L(m,n) &= \sum_{i,j=-\infty}^{\infty} \phi^R_{ij} K_L((m,n);\ (i,j)) \\
&= \frac{1}{4\pi^2}\iint_{\Omega\Omega} e^{i\omega m} e^{i\lambda n} \frac{\sigma(\omega,\lambda)}{\eta_L(\omega,\lambda)}\, \eta_L(\omega,\lambda)\, d\omega\, d\lambda \\
&= K_S(m,n).
\end{aligned}$$

It follows that Φ_R is a robust symmetric filter for the game $(\mathbb{F}(\mathbb{K}(\mathbb{U}')),\mathbb{U}',D)$.

4.6. Conclusion

In this chapter, we have investigated the application of reproducing kernel Hilbert space theory to the problems of robust signal detection and estimation. In particular, we have characterized minimax robust solutions for the general L^2-estimation problem in the presence

[5] Note that such a c always exists for the case in which $\sigma(\omega,\lambda) > 0$ for all $(\omega,\lambda) \in \Omega^2$.

of uncertainty regarding the second-order structure of the problem. Also, we have discussed conditions under which robust solutions to this problem are guaranteed to exist.

These results for the L^2-estimation problem are remarkably similar to results given by Poor in [53] relating to robust matched filtering. In order to more clearly reveal the similarities between the two problems, we have reformulated the robust matched filtering problem in an RKHS context. Within this context, we have seen that most of our results pertaining to robust L^2 estimation are also valid for the robust matched filtering problem. Many of these results can be seen to be extensions of those given in [53].

Finally, we have considered the problem of robust quadratic detection of Gaussian signals in the presence of Gaussian noise where the deflection ratio is used as a performance criterion. We have shown that this problem can also be formulated in an RKHS context. Using this formulation, we have shown that the robust quadratic detection problem is essentially analogous to the robust matched filtering problem. It should be mentioned that the deflection ratio may not be the best measure of performance for the robust quadratic detection problem. While it is well known that the deflection ratio possesses certain desirable properties in small-signal situations (see, for example, [73]), a better measure of performance in the general case may be the so-called *modified deflection ratio*. A discussion of this performance measure in relation to the deflection ratio and other generalized signal-to-noise ratios is given in appendix D. The robust quadratic detection problem incorporating the modified deflection ratio as a performance criterion can also be formulated and analyzed in an RKHS context.

The approach presented in this chapter, in addition to providing a unified view of the problems discussed above, provides a formulation that is useful for investigating robustness properties in other problems to which RKHS theory applies. For example, in a recent paper

[51], Picinbono and Duvaut discuss the design of optimal linear-quadratic detection and estimation strategies in non-Gaussian situations. It appears that this problem can be reformulated in RKHS terms for the purpose of designing robust linear-quadratic detectors. The structure and analysis of the problem would undoubtedly be very similar to that presented in Section 4.5.

Also, many different approaches to signal reconstruction and spectrum estimation are particular examples of a more general RKHS formulation (see, for example, [12] and [79]). When formulated in this general setting, these problems are seen to be analogous to the general L^2-estimation problem. This being the case, it would seem natural to apply the minimax techniques discussed in this chapter to the problem of signal reconstruction in the presence of noisy observations. The design of robust signal reconstruction and spectrum estimation procedures in this context is an interesting topic for further investigation.

CHAPTER 5

CONCLUSION

In this thesis, we have considered several different statistical signal processing problems, and we have applied reproducing kernel Hilbert space theory to the study of each. The thesis is not intended to be a study of the applications of RKHS theory; our primary goal has been to investigate the signal processing problems presented herein. Nevertheless, the work clearly demonstrates that RKHS techniques can be very useful and have a broad range of application.

The thesis is divided into two principal parts. In the first part (Chapter 3), we considered the problem of signal detection in fractional Gaussian noise. We were able to answer several interesting questions related to this problem; for example, we gave conditions that are necessary and sufficient for the problem to be nonsingular; we developed whitening filters, and we characterized the optimal detector in terms of the likelihood ratio. We have left unanswered, however, many equally interesting questions. For instance, we have not considered the problem of sequence detection in FGN, nor have we considered what mechanisms might be expected to give rise to additive FGN on communication channels. In short, there is much interesting research yet to be done with respect to this problem and with respect to other aspects of statistical signal processing in the presence of strongly dependent noise.

In second part of the thesis (Chapter 4), we studied some problems in robust detection and estimation. Applying RKHS techniques, we characterized minimax robust solutions for L^2 estimation, matched filtering, and quadratic detection in the presence of uncertainty regarding the relevant statistics. We also gave some results regarding the existence of solutions to these problems. The RKHS approach provided a general and unified framework in which to analyze these problems, and we believe that it can also be profitably applied to other problems, such as robust signal reconstruction.

APPENDIX A

SOME LEMMAS REFERENCED IN CHAPTER 3

In this appendix, we state and prove several technical lemmas that were referenced in Chapter 3. Throughout the appendix, μ will be the measure defined by (3.2.11), and Λ_H will be the subset of functions in $L^2(\mathbb{R})$ defined by (3.2.12).

Lemma A.1: If $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then $f \in \Lambda_H$.

Proof: Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. We must show that $\hat{f} \in L^2(\mu)$. To this end, recall that, since $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, it follows that $\hat{f} \in L^2(\mathbb{R})$, and, for all $\omega \in \mathbb{R}$,

$$\begin{aligned}
|\hat{f}(\omega)| &= \left| \int_{-\infty}^{\infty} e^{-i\omega t} f(t)\,dt \right| \\
&\le \int_{-\infty}^{\infty} |f(t)|\,dt \\
&\triangleq K < \infty.
\end{aligned}$$

Hence,

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 |\omega|^{1-2H}\,d\omega &\le \frac{1}{2\pi}\int_{-1}^{1} |\hat{f}(\omega)|^2 |\omega|^{1-2H}\,d\omega + \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}(\omega)|^2\,d\omega \\
&\le \frac{K^2}{2\pi}\int_{-1}^{1} |\omega|^{1-2H}\,d\omega + \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}(\omega)|^2\,d\omega \\
&< \infty.
\end{aligned}$$

Therefore, $\hat{f} \in L^2(\mu)$ and $f \in \Lambda_H$. ■

Lemma A.2: Let $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Then

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(-\omega)\,\overline{\hat{g}(-\omega)}\,|\omega|^{1-2H}\,d\omega = V_H H(2H-1)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\,\overline{g(s)}\,|t-s|^{2H-2}\,ds\,dt, \qquad \text{(A.1)}$$

where V_H is defined by (3.2.3).

Proof: Suppose first that f and g have compact support. Then it follows straightforwardly from Young's inequality (see, [15], page 232) that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(t)|\,|g(s)|\,|t-s|^{2H-2}\,ds\,dt < \infty.$$

Hence,

$$\begin{aligned}
\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(-\omega)\,\overline{\hat{g}(-\omega)}\,|\omega|^{1-2H}\,d\omega &= \lim_{N\to\infty}\frac{1}{2\pi}\int_{-N}^{N} \hat{f}(-\omega)\,\overline{\hat{g}(-\omega)}\,|\omega|^{1-2H}\,d\omega \\
&= \lim_{N\to\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\,\overline{g(s)}\left[\frac{1}{2\pi}\int_{-N}^{N} e^{i\omega(t-s)}|\omega|^{1-2H}\,d\omega\right]ds\,dt \\
&= \lim_{N\to\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\,\overline{g(s)}\left[\frac{1}{\pi}\int_{0}^{N} \cos(\omega|t-s|)\,\omega^{1-2H}\,d\omega\right]ds\,dt \\
&= \lim_{N\to\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\,\overline{g(s)}\,|t-s|^{2H-2}\left[\frac{1}{\pi}\int_{0}^{N|t-s|} \cos(\lambda)\lambda^{1-2H}\,d\lambda\right]ds\,dt.
\end{aligned}$$

Now,

$$\lim_{x\to\infty}\frac{1}{\pi}\int_0^x \cos(\lambda)\lambda^{1-2H}\,d\lambda = V_H H(2H-1),$$

and there exists a constant $K > 0$ such that, for all $x > 0$,

$$\left| \int_0^x \cos(\lambda)\lambda^{1-2H}\, d\lambda \right| < K.$$

Therefore, by the dominated convergence theorem,

$$\lim_{N\to\infty} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\overline{g(s)}\, |t-s|^{2H-2} \left[\frac{1}{\pi} \int_0^{N|t-s|} \cos(\lambda)\lambda^{1-2H}\, d\lambda \right] dsdt$$

$$= V_H H(2H-1) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\overline{g(s)}\, |t-s|^{2H-2} dsdt,$$

and (A.1) follows for f and g with compact support. To prove (A.1) for general $f, g \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, let $\{f_N\}_{N=1}^{\infty}$ and $\{g_N\}_{N=1}^{\infty}$ be defined by

$$f_N(t) \triangleq f(t) I_{[-N,N]}(t), \qquad t \in \mathbb{R},$$

and

$$g_N(t) \triangleq g(t) I_{[-N,N]}(t), \qquad t \in \mathbb{R}.$$

Clearly, $f_N \to f$ and $g_N \to g$ in $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$, and it follows from the proof of Lemma 3.2.1 that $\hat{f}_N \to \hat{f}$ and $\hat{g}_N \to \hat{g}$ in $L^2(\mu)$. Since f_N and g_N have compact support for all N, we get

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(-\omega)\overline{\hat{g}(-\omega)}\, |\omega|^{1-2H} d\omega = \lim_{N\to\infty} \int_{-\infty}^{\infty} \hat{f}_N(-\omega)\overline{\hat{g}_N(-\omega)}\, |\omega|^{1-2H} d\omega$$

$$= \lim_{N\to\infty} V_H H(2H-1) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_N(t)\overline{g_N(s)}\, |t-s|^{2H-2} dsdt$$

$$= \lim_{N\to\infty} V_H H(2H-1) \int_{-N}^{N}\int_{-N}^{N} f(t)\overline{g(s)}\, |t-s|^{2H-2} dsdt$$

$$= V_H H(2H-1) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)\overline{g(s)}\, |t-s|^{2H-2} dsdt. \quad \blacksquare$$

Lemma A.3: Let $\hat{S} \triangleq \{\frac{e^{i\omega t}-1}{i\omega};\ t \in \mathbb{R}\}$. Then $\hat{S}$ spans $L^2(\mu)$.

Proof: Clearly, $S \triangleq \{I_{[0,t]};\ t \in \mathbb{R}\}$ spans $L^2(\mathbb{R})$, and, since

$$\frac{e^{i\omega t}-1}{i\omega} = \hat{I}_{[0,t]}(-\omega), \qquad \forall\ \omega \in \mathbb{R},$$

it follows from the Fourier-Plancherel theorem that $\hat{S}$ spans $L^2(\mathbb{R})$. Hence, $\{\hat{g}_t(\omega) \triangleq \frac{e^{i\omega t}-1}{i\omega}|\omega|^{H-1/2}e^{i\,\mathrm{sgn}(\omega)(H-1/2)\frac{\pi}{2}};\ t \in \mathbb{R}\}$ spans $L^2(\mu)$ and it is sufficient to show that any $\hat{g}_t$ can be approximated in $L^2(\mu)$ by a finite linear combination of functions in $\hat{S}$. Now, it is straightforward to show that $\hat{g}_t \in L^2(\mathbb{R})$ with inverse Fourier transform $g_t \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ given by

$$g_t(\tau) = \frac{1}{\Gamma(3/2-H)}\left[I_{(0,\infty)}(\tau)(|t+\tau|^{1/2-H} - |\tau|^{1/2-H}) + I_{(-t,0]}(\tau)|t+\tau|^{1/2-H}\right], \qquad \tau \in \mathbb{R}.$$

Since $g_t \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, it is clear that there exists a sequence of functions $\{\phi_n(\tau) \triangleq \sum_{i=1}^{n} a_i I_{[0,t_i]}(-\tau)\}_{n=1}^{\infty}$ converging to g_t in $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$. It follows from the proof of Lemma 3.2.1 that $\hat{\phi}_n \to \hat{g}_t$ in $L^2(\mu)$, and, since

$$\hat{\phi}_n(\omega) = \sum_{i=1}^{n} a_i \frac{e^{i\omega t_i}-1}{i\omega}, \qquad \forall\ \omega \in \mathbb{R},$$

the result follows. ■

Lemma A.4: Let $[0,T]$ be a compact interval and, for $t \in [0,T]$, let f_t be given by

$$f_t(u) \triangleq \frac{1}{\Gamma(H-1/2)} I_{[0,t]}(u) u^{1/2-H} \int_u^t \tau^{H-1/2}(\tau-u)^{H-3/2} d\tau, \qquad u \in [0,T].$$

Then the set of functions $\{f_t;\ t \in [0,T]\}$ spans $L^2([0,T])$.

Proof: It suffices to show that the only function in $L^2([0,T])$ orthogonal to $\{f_t;\ t \in [0,T]\}$ is the zero function. To this end, let $g \in L^2([0,T])$ and suppose that, for all $t \in [0,T]$,

$$0 = \int_0^T f_t(u)\,\overline{g(u)}\,du$$
$$= \frac{1}{\Gamma(H-\frac{1}{2})} \int_0^t u^{\frac{1}{2}-H} \left[\int_u^t \tau^{H-\frac{1}{2}} (\tau-u)^{H-\frac{3}{2}} d\tau \right] \overline{g(u)}\,du$$
$$= \frac{1}{\Gamma(H-\frac{1}{2})} \int_0^t \tau^{H-\frac{1}{2}} \left[\int_0^\tau (\tau-u)^{H-\frac{3}{2}} u^{\frac{1}{2}-H}\, \overline{g(u)}\,du \right] d\tau.$$

Differentiating with respect to t, we see that, for almost all $t \in [0,T]$,

$$0 = \frac{1}{\Gamma(H-\frac{1}{2})} t^{H-\frac{1}{2}} \int_0^t (t-u)^{H-\frac{3}{2}} u^{\frac{1}{2}-H}\, \overline{g(u)}\,du.$$

Hence, for all $t \in [0,T]$,

$$0 = \frac{1}{\Gamma(\frac{3}{2}-H)} \int_0^t (t-s)^{\frac{1}{2}-H} \frac{1}{\Gamma(H-\frac{1}{2})} \left[\int_0^s (s-u)^{H-\frac{3}{2}} u^{\frac{1}{2}-H}\, \overline{g(u)}\,du \right] ds$$
$$= \frac{1}{\Gamma(\frac{3}{2}-H)\Gamma(H-\frac{1}{2})} \int_0^t u^{\frac{1}{2}-H}\, \overline{g(u)} \left[\int_u^t (t-s)^{\frac{1}{2}-H} (s-u)^{H-\frac{3}{2}} ds \right] du$$
$$= \int_0^t u^{\frac{1}{2}-H}\, \overline{g(u)}\,du.$$

Differentiating again with respect to t, we see that, for almost all $t \in [0,T]$,

$$0 = t^{\frac{1}{2}-H}\, \overline{g(t)}.$$

Hence, $g = 0$ almost everywhere in $[0,T]$, and it follows that $\{f_t;\ t \in [0,T]\}$ spans $L^2([0,T])$. ■

APPENDIX B

PROOF OF THEOREM 4.3.5

In this appendix, we prove Theorem 4.3.5. Throughout the appendix, I will represent an arbitrary index set, $\boldsymbol{U}$ an admissible uncertainty class defined on I, $\mathbb{K}(\boldsymbol{U})$ the class of covariance functions contained in $\boldsymbol{U}$, and $\mathbb{F}(\mathbb{K}(\boldsymbol{U}))$ the class of filters defined on $\mathbb{K}(\boldsymbol{U})$. In order to prove Theorem 4.3.5, we will need the following technical lemmas.

Lemma B.1: For any covariance function K defined on I and any $C > 0$, $\boldsymbol{H}(CK) = \boldsymbol{H}(K)$, and

$$\langle \cdot, * \rangle_{\boldsymbol{H}(CK)} = \frac{1}{C} \langle \cdot, * \rangle_{\boldsymbol{H}(K)}.$$

Further, if K_0 and K_1 are two covariance functions on I, and $K_1 \ll CK_0$, then $\boldsymbol{H}(K_1) \subseteq \boldsymbol{H}(K_0)$, and

$$\|f\|^2_{\boldsymbol{H}(K_0)} \leq C \|f\|^2_{\boldsymbol{H}(K_1)}, \qquad \forall f \in \boldsymbol{H}(K_1).$$

Proof: See [1], §I.7 and §I.13. ■

Lemma B.2: If K_0 and K_1 are covariance functions on I and $K_1 \ll CK_0$, then $\mathbf{h} \in \mathbb{F}(K_0)$ implies $\mathbf{h} \in \mathbb{F}(K_1)$.

Proof: Let $\mathbf{h} = \{\bar{\mathbf{h}}_N\}_{N=1}^{\infty}$, where $\bar{\mathbf{h}}_1, \bar{\mathbf{h}}_2, \cdots$ are finite filters. Then, since $CK_0 - K_1$ is nonnegative definite on I and $\mathbf{h} \in \mathbb{F}(K_0)$,

$$
\begin{aligned}
0 &\le \lim_{M,N\to\infty} \|\bar{\mathbf{h}}_M K_1 - \bar{\mathbf{h}}_N K_1\|^2_{\boldsymbol{H}(K_1)} \\
&= \lim_{M,N\to\infty} \left\langle (\bar{\mathbf{h}}_M - \bar{\mathbf{h}}_N) K_1, (\bar{\mathbf{h}}_M - \bar{\mathbf{h}}_N) \right\rangle \\
&\le \lim_{M,N\to\infty} C \left\langle (\bar{\mathbf{h}}_M - \bar{\mathbf{h}}_N) K_0, (\bar{\mathbf{h}}_M - \bar{\mathbf{h}}_N) \right\rangle \\
&= \lim_{M,N\to\infty} C \|\bar{\mathbf{h}}_M K_0 - \bar{\mathbf{h}}_N K_0\|^2_{\boldsymbol{H}(K_0)} \\
&= 0.
\end{aligned}
$$

It follows that $\mathbf{h} \in \boldsymbol{F}(K_1)$. ∎

Lemma B.3: Suppose $K_1 \ll CK_0$ and let $K_\alpha = (1-\alpha)K_0 + \alpha K_1$ for $0 < \alpha < 1$. Then $\boldsymbol{H}(K_\alpha) = \boldsymbol{H}(K_0)$, and, for all α sufficiently small,

$$(1-\alpha)K_0 \ll K_\alpha \ll (1-\alpha+\sqrt{\alpha})K_0. \tag{B.1}$$

Hence, by Lemma B.1,

$$\frac{1}{1-\alpha+\sqrt{\alpha}}\|\cdot\|^2_{\boldsymbol{H}(K_0)} \le \|\cdot\|^2_{\boldsymbol{H}(K_\alpha)} \le \frac{1}{1-\alpha}\|\cdot\|^2_{\boldsymbol{H}(K_0)}, \tag{B.2}$$

and $\|\cdot\|^2_{\boldsymbol{H}(K_\alpha)} \to \|\cdot\|^2_{\boldsymbol{H}(K_0)}$ as $\alpha \to 0$.

Proof: Clearly, $(1-\alpha)K_0 \ll K_\alpha \ll (C+1)K_0$. This establishes the left-hand side of (B.1) and implies (by Lemma B.1) that $\boldsymbol{H}(K_\alpha) = \boldsymbol{H}(K_0)$ for all $0 < \alpha < 1$. Also,

$$
\begin{aligned}
(1-\alpha+\sqrt{\alpha})K_0 - K_\alpha &= (1-\alpha+\sqrt{\alpha})K_0 - (1-\alpha)K_0 - \alpha K_1 \\
&= \sqrt{\alpha}K_0 - \alpha K_1 \\
&= \alpha\left[\frac{1}{\sqrt{\alpha}}K_0 - K_1\right].
\end{aligned}
$$

Hence, for all $0 < \alpha < C^{-2}$, $K_\alpha \ll (1-\alpha+\sqrt{\alpha})K_0$, which establishes the right-hand side of (B.1) and proves the lemma. ∎

Lemma B.4: Let K_0 and K_1 be covariance functions such that $K_1 \ll CK_0$. Let $s_0 \in \boldsymbol{H}(K_0)$ and $K_\alpha = (1-\alpha)K_0 + \alpha K_1$ for $0 < \alpha < 1$. Choose $\mathbf{h}_0 \in \boldsymbol{F}(K_0)$ such that $\mathbf{h}_0 K_0 = s_0$ and $\mathbf{h}_\alpha \in \boldsymbol{F}(K_\alpha)$ such that $\mathbf{h}_\alpha K_\alpha = s_0$ (recall $\boldsymbol{H}(K_0) = \boldsymbol{H}(K_\alpha)$, by Lemma B.3). Then

$$\lim_{\alpha\to 0} \left\langle (\mathbf{h}_\alpha - \mathbf{h}_0)K_0, (\mathbf{h}_\alpha - \mathbf{h}_0) \right\rangle = \lim_{\alpha \to 0} \|\mathbf{h}_\alpha K_0 - \mathbf{h}_0 K_0\|^2_{\boldsymbol{H}(K_0)} = 0, \tag{B.3}$$

and

$$\lim_{\alpha\to 0} \left\langle (\mathbf{h}_\alpha - \mathbf{h}_0)K_1, (\mathbf{h}_\alpha - \mathbf{h}_0) \right\rangle = \lim_{\alpha \to 0} \|\mathbf{h}_\alpha K_1 - \mathbf{h}_0 K_1\|^2_{\boldsymbol{H}(K_1)} = 0. \tag{B.4}$$

Proof: Notice first that (by Lemmas B.1 and B.3) $\boldsymbol{H}(K_1) \subseteq \boldsymbol{H}(K_0) = \boldsymbol{H}(K_\alpha)$. Let $\mathbf{h}_0 = \{\bar{\mathbf{h}}_N^0\}_{N=1}^\infty$. Then

$$\begin{aligned}
0 &\le \lim_{N\to\infty} \|s_0 - \bar{\mathbf{h}}_N^0 K_0\|^2_{\boldsymbol{H}(K_\alpha)} \\
&\le \lim_{N\to\infty} \frac{1}{1-\alpha} \|s_0 - \bar{\mathbf{h}}_N^0 K_0\|^2_{\boldsymbol{H}(K_0)} \qquad \text{(by (B.2))} \\
&= \frac{1}{1-\alpha} \|s_0 - \mathbf{h}_0 K_0\|^2_{\boldsymbol{H}(K_0)} \\
&= 0.
\end{aligned}$$

So $\mathbf{h}_0 K_0$ converges to s_0 in $\boldsymbol{H}(K_\alpha)$ as well as in $\boldsymbol{H}(K_0)$. Further, by Lemma B.2, $\boldsymbol{F}(K_\alpha) = \boldsymbol{F}(K_0) \subseteq \boldsymbol{F}(K_1)$, and it follows from an argument similar to the above that $\mathbf{h}_0 K_1$ and $\mathbf{h}_\alpha K_1$ both converge in $\boldsymbol{H}(K_0)$ and $\boldsymbol{H}(K_\alpha)$ as well as in $\boldsymbol{H}(K_1)$. Hence,

$$\left\langle (\mathbf{h}_\alpha-\mathbf{h}_0)K_0,\ (\mathbf{h}_\alpha-\mathbf{h}_0)\right\rangle \le \frac{1}{1-\alpha}\left\langle (\mathbf{h}_\alpha-\mathbf{h}_0)K_\alpha,\ (\mathbf{h}_\alpha-\mathbf{h}_0)\right\rangle$$
$$= \frac{1}{1-\alpha}\|\mathbf{h}_\alpha K_\alpha - \mathbf{h}_0 K_\alpha\|^2_{\mathbb{H}(K_\alpha)}$$
$$= \frac{1}{1-\alpha}\|\mathbf{h}_\alpha K_\alpha - (1-\alpha)\mathbf{h}_0 K_0 - \alpha\mathbf{h}_0 K_1\|^2_{\mathbb{H}(K_\alpha)}$$
$$= \frac{1}{1-\alpha}\|s_0 - (1-\alpha)s_0 - \alpha\mathbf{h}_0 K_1\|^2_{\mathbb{H}(K_\alpha)}$$
$$= \frac{\alpha^2}{1-\alpha}\|s_0 - \mathbf{h}_0 K_1\|^2_{\mathbb{H}(K_\alpha)}.$$

(B.3) follows by letting $\alpha \to 0$ and noting that $\|s_0 - \mathbf{h}_0 K_1\|^2_{\mathbb{H}(K_\alpha)}$ remains bounded (by lemma B.3). Finally, since $K_1 \ll CK_0$,

$$0 \le \left\langle (\mathbf{h}_\alpha-\mathbf{h}_0)K_1,\ (\mathbf{h}_\alpha-\mathbf{h}_0)\right\rangle \le C\left\langle (\mathbf{h}_\alpha-\mathbf{h}_0)K_0,\ (\mathbf{h}_\alpha-\mathbf{h}_0)\right\rangle,$$

so (B.4) follows from (B.3). ■

Proof of Theorem 4.3.5: It follows immediately from lemma B.2 that $\mathbf{h}_R \in \mathbb{F}(\mathbb{K}(\mathbb{U}))$. By Lemma 4.3.2, $M^*(s,K) = \sigma_Z^2 - \langle s, s\rangle_{\mathbb{H}(K)}$ for $(s,K) \in \mathbb{U}$, and by Lemma 4.3.3, M^* is concave on $\mathbb{U}$. Therefore,

$$\langle s_L, s_L\rangle_{\mathbb{H}(K_L)} \le \langle s, s\rangle_{\mathbb{H}(K)}, \qquad \forall\ (s,K) \in \mathbb{U},$$

if and only if, for all $(s,K) \in \mathbb{U}$,

$$\lim_{\alpha\to 0}\frac{1}{\alpha}\left[\langle s_\alpha, s_\alpha\rangle_{\mathbb{H}(K_\alpha)} - \langle s_L, s_L\rangle_{\mathbb{H}(K_L)}\right] \ge 0,$$

where $s_\alpha = (1-\alpha)s_L + \alpha s$ and $K_\alpha = (1-\alpha)K_L + \alpha K$. Separating terms and applying Lemma B.3, we get

$$\lim_{\alpha\to 0}\frac{1}{\alpha}\left[\langle s_\alpha, s_\alpha\rangle_{\boldsymbol{H}(K_\alpha)} - \langle s_L, s_L\rangle_{\boldsymbol{H}(K_L)}\right] = 2\langle s, s_L\rangle_{\boldsymbol{H}(K_L)} - 2\langle s_L, s_L\rangle_{\boldsymbol{H}(K_L)}$$

$$+ \lim_{\alpha\to 0}\frac{1}{\alpha}\left[\langle s_L, s_L\rangle_{\boldsymbol{H}(K_\alpha)} - \langle s_L, s_L\rangle_{\boldsymbol{H}(K_L)}\right]$$

$$= 2\langle s, \mathbf{h}_R\rangle - 2\langle s_L, \mathbf{h}_R\rangle_{\boldsymbol{H}(K_L)}$$

$$+ \lim_{\alpha\to 0}\frac{1}{\alpha}\left[\langle s_L, s_L\rangle_{\boldsymbol{H}(K_\alpha)} - \langle s_L, s_L\rangle_{\boldsymbol{H}(K_L)}\right].$$

Now, as in the proof of Lemma B.4, we find $\mathbf{h}_\alpha \in \boldsymbol{F}(\boldsymbol{K}(\boldsymbol{U}))$ such that $\mathbf{h}_\alpha K_\alpha = s_L$ with convergence in both $\boldsymbol{H}(K_\alpha)$ and $\boldsymbol{H}(K_L)$. Then

$$\langle s_L, s_L\rangle_{\boldsymbol{H}(K_L)} = \langle \mathbf{h}_\alpha K_\alpha, \mathbf{h}_\alpha K_\alpha\rangle_{\boldsymbol{H}(K_L)}$$

$$= (1-\alpha)^2\langle \mathbf{h}_\alpha K_L, \mathbf{h}_\alpha\rangle + 2\alpha(1-\alpha)\langle \mathbf{h}_\alpha K, \mathbf{h}_\alpha\rangle + \alpha^2\|\mathbf{h}_\alpha K\|^2_{\boldsymbol{H}(K_L)},$$

and

$$\langle s_L, s_L\rangle_{\boldsymbol{H}(K_\alpha)} = \langle \mathbf{h}_\alpha K_\alpha, \mathbf{h}_\alpha\rangle$$

$$= (1-\alpha)\langle \mathbf{h}_\alpha K_L, \mathbf{h}_\alpha\rangle + \alpha\langle \mathbf{h}_\alpha K, \mathbf{h}_\alpha\rangle.$$

Hence,

$$\lim_{\alpha\to 0}\frac{1}{\alpha}\left[\langle s_L, s_L\rangle_{\boldsymbol{H}(K_\alpha)} - \langle s_L, s_L\rangle_{\boldsymbol{H}(K_L)}\right] = \lim_{\alpha\to 0}\left[(1-\alpha)\langle \mathbf{h}_\alpha K_L, \mathbf{h}_\alpha\rangle + (2\alpha-1)\langle \mathbf{h}_\alpha K, \mathbf{h}_\alpha\rangle\right.$$

$$\left. - \alpha\|\mathbf{h}_\alpha K\|^2_{\boldsymbol{H}(K_L)}\right]$$

$$= \langle \mathbf{h}_R K_L, \mathbf{h}_R\rangle - \langle \mathbf{h}_R K, \mathbf{h}_R\rangle \qquad \text{(by Lemma B.4).}$$

$$= \langle s_L, \mathbf{h}_R\rangle - \langle \mathbf{h}_R K, \mathbf{h}_R\rangle.$$

Therefore,

$$\lim_{\alpha\to 0}\frac{1}{\alpha}\left[\langle s_\alpha, s_\alpha\rangle_{\boldsymbol{H}(K_\alpha)} - \langle s_L, s_L\rangle_{\boldsymbol{H}(K_L)}\right] = 2\langle s, \mathbf{h}_R\rangle - \langle s_L, \mathbf{h}_R\rangle - \langle \mathbf{h}_R K, \mathbf{h}_R\rangle.$$

It follows that (4.3.9) holds if and only if (4.3.6) holds. This proves the theorem. ■

APPENDIX C

RKHS APPROXIMATION LEMMA AND PROOF OF LEMMA 4.4.8

RKHS Approximation Lemma: Let I be a separable metric space and let $\{I_N\}_{N=1}^{\infty}$ be a sequence of subsets of I that is monotone increasing $(I_1 \subset I_2 \subset \cdots)$ and such that $\bigcup_{N=1}^{\infty} I_N$ is dense in I. Let K be a continuous covariance function on I, and let K_N be the restriction of K to I_N. Let s be a continuous function on I and let s_N be its restriction to I_N. If $s \in \boldsymbol{H}(K)$, then $s_N \in \boldsymbol{H}(K_N)$ for all N and

$$\|s\|_{\boldsymbol{H}(K)} = \lim_{N \to \infty} \|s_N\|_{\boldsymbol{H}(K_N)}.$$

Conversely, if $s_N \in \boldsymbol{H}(K_N)$ for all N, then

$$\lim_{N \to \infty} \|s_N\|_{\boldsymbol{H}(K_N)} < \infty$$

only if $s \in \boldsymbol{H}(K)$.

Proof: See [47], pp. 316-319. ■

Proof of Lemma 4.4.8: Suppose that $(h_R; (s_L, K_L))$ is not a saddle point for $(\boldsymbol{F}(\boldsymbol{U}), \boldsymbol{U}, \Delta)$. Then there exists $(s, K) \in \boldsymbol{U}$ such that

$$\Delta(h_R; (s,K)) = \frac{\left[\langle h_R, s \rangle_{\boldsymbol{H}_0}\right]^2}{\langle h_R, K h_R \rangle_{\boldsymbol{H}_0}}$$
$$< \frac{\left[\langle h_R, s_L \rangle_{\boldsymbol{H}_0}\right]^2}{\langle h_R, K_L h_R \rangle_{\boldsymbol{H}_0}} = \Delta(h_R; (s_L, K_L)).$$

Choose $0 < \alpha < 1$ and let $s_\alpha = (1-\alpha)s + \alpha s_L$ and $K_\alpha = (1-\alpha)K + \alpha K_L$. It follows (see the proof of Lemma 4.4.1 or [53], Property 2) that

$$\Delta(h_R;\,(s_\alpha,K_\alpha)) \le (1-\alpha)\Delta(h_R;\,(s,K)) + \alpha\Delta(h_R;\,(s_L,K_L))$$
$$< \Delta(h_R;\,(s_L,K_L)).$$

Now, since $S \subset \boldsymbol{H}_0$ is convex with a nonempty interior, it is easily established that the interior points of S are dense in S. Further, since K_α is strictly positive, the range of K_α is dense in $\boldsymbol{H}_0$. These facts imply that there exists a sequence $\{s_n\}_{n=1}^{\infty} \subset S \cap \boldsymbol{H}(K_\alpha)$ such that $\|s_n - s_\alpha\|^2_{\boldsymbol{H}_0} \to 0$. It follows that $\langle h_R, s_n\rangle_{\boldsymbol{H}_0} \to \langle h_R, s_\alpha\rangle_{\boldsymbol{H}_0}$, and, since $\langle h_R, K_\alpha h_R\rangle_{\boldsymbol{H}_0} > 0$, $\Delta(h_R;\,(s_n,K_\alpha)) \to \Delta(h_R;\,(s_\alpha,K_\alpha))$. Hence, there exists an s_N such that $(s_N,K_\alpha) \in \boldsymbol{U}'$ and

$$\Delta(h_R;\,(s_N,K_\alpha)) < \Delta(h_R;\,(s_L,K_L)).$$

This contradicts the hypothesis that $(h_R;\,(s_L,K_L))$ is a saddle point for $(\mathbb{F}(\mathbb{K}(\boldsymbol{U}')),\boldsymbol{U}',\Delta)$. ■

APPENDIX D

GENERALIZED SIGNAL-TO-NOISE RATIOS IN QUADRATIC DETECTION

D.1. Introduction

In this appendix, we discuss some properties of the so-called generalized signal-to-noise ratios with regard to their use in evaluating the performance of quadratic detectors used to discriminate between two Gaussian hypotheses. A somewhat more detailed discussion is given in [6]. Throughout the appendix, $\mathbf{X} \triangleq \{X(t), t \in I\}$ will represent an observed process defined on some index set I, which is assumed to be a separable metric space. We assume that $\mathbf{X}$ is a zero-mean, Gaussian process with one of two continuous covariance functions; the object being to test the hypotheses:

$$\begin{aligned} &H_0\text{: } \mathbf{X} \text{ has covariance function } K_0 \\ \text{versus} \qquad & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(D.1.1)} \\ &H_1\text{: } \mathbf{X} \text{ has covariance function } K_1, \end{aligned}$$

A special case of considerable practical importance is the signal detection problem, in which H_1 corresponds to a Gaussian signal in additive, independent, Gaussian noise, and H_0 corresponds to noise only. In keeping with this example, we define

$$K_S \triangleq K_1 - K_0;$$

however, one should keep in mind that, in general, K_S need not be a covariance function. We assume also that Problem (D.1.1) is nonsingular.

It is common practice to decide Problem (D.1.1) by using a quadratic detector; that is, a detector in which the test statistic $\Phi(X)$ is a quadratic functional of the observed process. Indeed, the optimal detectors[6] generally take this form. Because it is difficult to evaluate the probabilities of error for such a detector, other measures of performance are often used. Among these are the *generalized signal-to-noise ratios* (GSNR's), which we will represent as a parametric family $\{D_\alpha, 0 \le \alpha \le 1\}$. Given any quadratic statistic $\Phi(X)$, $D_\alpha(\Phi)$ is defined as [17]

$$D_\alpha(\Phi) \triangleq \frac{\left[E_1\{\Phi(X)\} - E_0\{\Phi(X)\}\right]^2}{(1-\alpha)\mathrm{Var}_0\{\Phi(X)\} + \alpha \mathrm{Var}_1\{\Phi(X)\}},$$

where the subscripts 0 and 1 indicate expectations taken under H_0 and H_1, respectively. D_0 is referred to as the *deflection ratio*, which we discussed in Section 4.5, and D_1 as the *complementary deflection ratio*. We will show that every performance measure in this class can be represented as a ratio of inner products in a reproducing kernel Hilbert space. Such a representation has several advantages. Apart from being mathematically appealing, it leads straightforwardly to a characterization of the quadratic statistic $\Phi_\alpha^*(X)$ that maximizes $D_\alpha(\Phi)$ for each $\alpha \in [0,1]$. Also, it clearly reveals one of the weaknesses of the GSNR as a measure of performance for quadratic detectors and leads naturally to consideration of an alternative measure, which is superior to the GSNR in some respects.

[6]Throughout this appendix, the term "optimal detector" refers to a detector in which the test statistic $\Phi(X)$ is (with probability one) a monotone function of the likelihood ratio $L(X)$.

D.2. GSNR Representation

As in Section 4.5, we define a quadratic test statistic as any random variable $\Phi(X)$ that can be written in the form

$$\Phi(X) = \lim_{N \to \infty} \tilde{\Phi}_N(X) \triangleq \lim_{N \to \infty} \sum_{i,j=1}^{\infty} \phi_{ij}(N) X(t_i) X(t_j), \tag{D.2.1}$$

where $\{t_i\}_{i=1}^{\infty} \subseteq I$ is a sequence of observation points (*depending on* $\Phi(X)$), and, for each positive integer N, the infinite-dimensional, real matrix $(\phi_{ij}(N))$ is symmetric and has only finitely many nonzero elements. The limit is taken in the mean-square sense; i.e., we assume that, for $n = 0$ and 1,

$$\lim_{N \to \infty} \mathrm{E}_n [\Phi(X) - \tilde{\Phi}_N(X)]^2 = 0. \tag{D.2.2}$$

Now, let $\alpha \in [0,1]$ and define K_α: $I^2 \times I^2 \to \boldsymbol{R}$ by

$$K_\alpha((t_1,\tau_1); (t_2,\tau_2)) \triangleq (1-\alpha) K_0(t_1,t_2) K_0(\tau_1,\tau_2) + \alpha K_1(t_1,t_2) K_1(\tau_1,\tau_2).$$

Clearly, each K_α is symmetric and nonnegative definite on $I^2 \times I^2$. Hence, for each $\alpha \in [0,1]$, there exists an RKHS $\boldsymbol{H}(K_\alpha)$ with reproducing kernel K_α. If we let $\mathbb{K} \triangleq \{K_\alpha: 0 \le \alpha \le 1\}$, then it follows from (D.2.1) and (D.2.2) that, corresponding to each quadratic statistic $\Phi(X)$, there exists a symmetric filter $\Phi \in \mathbb{F}(\mathbb{K})$, and (see [3] for details)

$$
\begin{aligned}
D_\alpha(\Phi) &\triangleq \frac{\left[E_1\{\Phi(X)\} - E_0\{\Phi(X)\}\right]^2}{(1-\alpha)\mathrm{Var}_0\{\Phi(X)\} + \alpha \mathrm{Var}_1\{\Phi(X)\}} \\
&= \frac{\left[\lim\limits_{N\to\infty} \sum\limits_{i,j=1}^{\infty} \phi_{ij}(N) K_S(t_j,t_i)\right]^2}{2\left[\lim\limits_{N\to\infty} \sum\limits_{i,j=1}^{\infty} \sum\limits_{k,l=1}^{\infty} \phi_{ij}(N)\phi_{kl}(N)[(1-\alpha)K_0(t_k,t_i)K_0(t_l,t_j) + \alpha K_1(t_k,t_i)K_1(t_l,t_j)]\right]} \\
&= \frac{\left[\lim\limits_{N\to\infty} \sum\limits_{i,j=1}^{\infty} \phi_{ij}(N) K_S(t_j,t_i)\right]^2}{2\left[\lim\limits_{N\to\infty} \sum\limits_{i,j=1}^{\infty} \sum\limits_{k,l=1}^{\infty} \phi_{ij}(N)\phi_{kl}(N) K_\alpha((t_k,t_l);\ (t_i,t_j))\right]} \qquad \text{(D.2.3)} \\
&= \frac{\left[\langle K_S,\ \Phi K_\alpha \rangle_{\boldsymbol{H}(K_\alpha)}\right]^2}{2\langle \Phi K_\alpha,\ \Phi K_\alpha \rangle_{\boldsymbol{H}(K_\alpha)}}.
\end{aligned}
$$

Note that in order to express D_α in the final form given in Equation (D.2.3), we have used the fact that $K_S \in \boldsymbol{H}(K_\alpha)$, which follows from the nonsingularity conditions for Problem (D.1.1), as stated in Theorem 2.1.4.

One of the obvious advantages of representing D_α in the form given by (D.2.3) is that it leads immediately to a characterization of the quadratic statistic $\Phi_\alpha^*(X)$ that maximizes $D_\alpha(\Phi)$ for any given α. It appears that such a characterization was not previously known [67]. To characterize $\Phi_\alpha^*(X)$, notice that (D.2.3) and the Schwarz inequality together imply that

$$D_\alpha(\Phi) \le \tfrac{1}{2} \langle K_S,\ K_S \rangle_{\boldsymbol{H}(K_\alpha)},$$

with equality if and only if $\Phi K_\alpha = cK_S$ for some $c \neq 0$. That is,

$$\Phi_\alpha^*(\mathrm{X}) = \lim_{N\to\infty} \tilde{\Phi}_N^*(\mathrm{X}) \triangleq \lim_{N\to\infty} \sum_{i,j=1}^{\infty} \phi_{ij}^*(N) X(t_i^*) X(t_j^*), \tag{D.2.4}$$

maximizes $D_\alpha(\Phi)$ if and only if there exists $c \neq 0$, such that

$$\begin{aligned} cK_S(t,s) &= \Phi_\alpha^* K_\alpha(t,s) \\ &\triangleq \lim_{N\to\infty} \tilde{\Phi}_N^* K_\alpha(t,s) \\ &= \lim_{N\to\infty} \sum_{i,j=1}^{\infty} \phi_{ij}^*(N) K_\alpha((t,s);\ (t_i^*,t_j^*)), \end{aligned} \tag{D.2.5}$$

where the limit in (D.2.5) is taken in $\boldsymbol{H}(K_\alpha)$ (which implies pointwise convergence, as well). Note that, even though we can always find $\Phi_\alpha^* \triangleq \{\tilde{\Phi}_N^*\}_{N=1}^{\infty}$ satisfying (D.2.5), the corresponding $\{\tilde{\Phi}_N^*(\mathrm{X})\}_{N=1}^{\infty}$ may not converge in mean-square, so we cannot guarantee that $\Phi_\alpha^*(\mathrm{X})$ exists. However, we do have

$$D_\alpha^* \triangleq \sup_{\Phi\in \boldsymbol{F}(K_\alpha)} D_\alpha(\Phi) = \tfrac{1}{2} \langle K_S, K_S \rangle_{\boldsymbol{H}(K_\alpha)}.$$

As a concrete example, let us consider the following simple problem.

Example D.2.1: Let $I = \{1, 2, \cdots, N\}$, and assume that K_0 and K_1 are positive definite $N \times N$ matrices. We are looking for a $N \times N$ symmetric matrix Φ_α^* such that the quadratic form

$$\Phi_\alpha^*(\mathrm{X}) \triangleq \mathrm{X}^T \Phi_\alpha^* \mathrm{X}$$

maximizes the GSNR

$$D_\alpha(\Phi) = \frac{[\mathrm{Tr}\ \Phi K_S]^2}{2[(1-\alpha)(\mathrm{Tr}\ K_0\Phi K_0\Phi) + \alpha(\mathrm{Tr}\ K_1\Phi K_1\Phi)]}$$

for fixed $\alpha \in [0,1]$. It follows from (D.2.5) that Φ_α^* must satisfy

$$cK_S = (1-\alpha)K_0\Phi_\alpha^* K_0 + \alpha K_1 \Phi_\alpha^* K_1 \tag{D.2.6}$$

for some $c \neq 0$. One solution to (D.2.6), for $c = 1$, is given by

$$\begin{aligned}\Phi_\alpha^* &= K_1^{-1}K_S[(1-\alpha)K_0K_1^{-1}K_0 + \alpha K_1]^{-1} \\ &= [(1-\alpha)K_0K_1^{-1}K_0 + \alpha K_1]^{-1} - K_1^{-1}K_0[(1-\alpha)K_0K_1^{-1}K_0 + \alpha K_1]^{-1}.\end{aligned} \tag{D.2.7}$$

This solution was established by diagonalizing K_0 and K_1 simultaneously, but it is easy to verify directly that it satisfies Equation (D.2.6). For $\alpha \neq 0$, we use a matrix inversion lemma ([30], page 19) to write

$$\begin{aligned}P &\triangleq [(1-\alpha)K_0K_1^{-1}K_0 + \alpha K_1]^{-1} \\ &= \frac{1}{\alpha}K_1^{-1} - \frac{1}{\alpha}(1-\alpha)K_1^{-1}K_0[(1-\alpha)K_0K_1^{-1}K_0 + \alpha K_1]^{-1}K_0K_1^{-1} \\ &= \frac{1}{\alpha}K_1^{-1} - \frac{1}{\alpha}(1-\alpha)K_1^{-1}K_0PK_0K_1^{-1}.\end{aligned}$$

Then, for Φ_α^* defined by Equation (D.2.7), we have

$$\begin{aligned}(1-\alpha)K_0\Phi_\alpha^*K_0 + \alpha K_1\Phi_\alpha^*K_1 &= (1-\alpha)K_0(P - K_1^{-1}K_0P)K_0 + \alpha K_1(P - K_1^{-1}K_0P)K_1 \\ &= (1-\alpha)K_0PK_0 - (1-\alpha)K_0K_1^{-1}K_0PK_0 + \alpha K_1PK_1 - \alpha K_0PK_1 \\ &= (1-\alpha)K_0PK_0 - (1-\alpha)K_0K_1^{-1}K_0PK_0 + K_1 - (1-\alpha)K_0PK_0 - K_0 \\ &\quad + (1-\alpha)K_0K_1^{-1}K_0PK_0 \\ &= K_1 - K_0 \\ &= K_S,\end{aligned}$$

as claimed. Verification for $\alpha = 0$ is straightforward.

For this example, it is easy to check that

$$D_\alpha^* = D_\alpha(\Phi_\alpha^*) = \tfrac{1}{2}\mathrm{Tr}\ \Phi_\alpha^* K_S.$$

The foregoing discussion reveals one of the weaknesses of the GSNR, to wit: under mild regularity conditions on Problem (D.1.1), there exists an optimal quadratic detector incorporating a test statistic $\Phi_{opt}(\mathbf{X})$, which does not generally maximize $D_\alpha(\Phi)$ for any $\alpha \in [0,1]$. To be more specific, in [23], Kailath and Weinert have shown that if Problem (D.1.1) is *strongly nonsingular*[7] (as defined in [23]), then there exists an optimal quadratic detector for (D.1.1), and $\Phi_{opt}(\mathbf{X})$ satisfies

$$\Phi_{opt}(\mathbf{X}) = \lim_{N \to \infty} \sum_{i,j=1}^{\infty} \phi_{ij}^{opt}(N) X(t_i^{opt}) X(t_j^{opt}), \qquad \text{(D.2.8)}$$

where

$$K_S(t,s) = \lim_{N \to \infty} \sum_{i,j=1}^{\infty} \phi_{ij}^{opt}(N) K_0(t, t_i^{opt}) K_1(s, t_j^{opt}). \qquad \text{(D.2.9)}$$

However, it follows from (D.2.5) that $\Phi_{opt}(\mathbf{X})$ maximizes D_α if and only if there exists $c \neq 0$ such that

$$cK_S(t,s) = \lim_{N \to \infty} \sum_{i,j=1}^{\infty} \phi_{ij}^{opt}(N) K_\alpha((t,s);\ (t_i^{opt}, t_j^{opt})). \qquad \text{(D.2.10)}$$

As a general rule, (D.2.9) and (D.2.10) will not be satisfied simultaneously, and we will be able to find a nonoptimal quadratic statistic $\Phi(\mathbf{X})$ such that

$$D_\alpha(\Phi) > D_\alpha(\Phi_{opt}).$$

For example, consider the following simple signal detection problem.

[7]Actually, by changing our definition of quadratic statistic slightly to avoid convergence problems, we could drop this restriction. The interested reader is referred to [23] or [57] for details. In any case, this is only an issue when I is an infinite set since (D.1.1) is always strongly nonsingular (trivially) if I is finite.

Example D.2.2: Let $I = \{1, 2, \cdots, N\}$, $K_0 = K_N$, and $K_1 = K_N + K_S$, where K_N and K_S are both $N \times N$ covariance matrices. Assuming K_N is positive definite, there exists an optimal quadratic detector with

$$\Phi_{opt}(X) \triangleq X^T \Phi_{opt}(X),$$

where the matrix Φ_{opt} is given by

$$\Phi_{opt} \triangleq K_0^{-1} K_S K_1^{-1} = K_0^{-1} - K_1^{-1}.$$

In order for $\Phi_{opt}(X)$ to maximize D_α, we must be able to find $c \neq 0$ such that

$$\begin{aligned} cK_S &= (1-\alpha)K_0 \Phi_{opt} K_0 + \alpha K_1 \Phi_{opt} K_1 \\ &= (1-\alpha)K_0(K_0^{-1} - K_1^{-1})K_0 + \alpha K_1(K_0^{-1} - K_1^{-1})K_1, \end{aligned}$$

which is possible if and only if

$$\alpha(K_N^{-1}K_S)^3 + (2\alpha - c)(K_N^{-1}K_S)^2 + (1-c)K_N^{-1}K_S = 0. \tag{D.2.11}$$

In particular, if the noise is white, so that $K_N^{-1}K_S = K_S$, then (D.2.11) is satisfied if and only if all of the nonzero eigenvalues of K_S satisfy

$$\alpha\lambda^2 + (2\alpha - c)\lambda + (1-c) = 0.$$

For this example, we already know that Φ_α^* defined by Equation (D.2.7) satisfies Equation (D.2.6) and maximizes D_α. Hence, if (D.2.11) is not satisfied, it follows that Φ_α^* is nonoptimal ($\Phi_\alpha^* \neq c\Phi_{opt}$ for any $c \neq 0$), and yet $D_\alpha(\Phi_\alpha^*) > D_\alpha(\Phi_{opt})$.

D.3. Alternative Performance Measure

Since it seems reasonable to want a performance measure for quadratic detectors that is always maximized by an optimal quadratic detector, one is led to search for alternatives to the

GSNR. In light of Equations (D.2.8) and (D.2.9) vis-à-vis Equations (D.2.3) through (D.2.5) a natural choice is

$$D_\mu(\Phi) \triangleq \frac{\left[\langle \Phi K_{01}, K_S \rangle_{H(K_0) \otimes H(K_1)}\right]^2}{2\,\langle \Phi K_{01}, \Phi K_{01} \rangle_{H(K_0) \otimes H(K_1)}}, \tag{D.3.1}$$

where $K_{01}: I^2 \times I^2 \to \mathbb{R}$ is defined by

$$K_{01}((t_1,\tau_1);\ (t_2,\tau_2)) \triangleq K_0(t_1,t_2)K_1(\tau_1,\tau_2).$$

The fact that $D_\mu(\Phi)$ is well defined for any quadratic statistic $\Phi(X)$ (i.e., $K_S \in H(K_{01}) = H(K_0) \otimes H(K_1)$ and $\Phi \in \mathbb{F}(K_{01})$) follows easily from the mean-square convergence in (D.2.2) and the nonsingularity conditions given in Theorem 2.1.4.

Equation (D.3.1) and the Schwarz inequality together imply that

$$D_\mu(\Phi) \le \tfrac{1}{2} \langle K_S, K_S \rangle_{H(K_0) \otimes H(K_1)},$$

with equality if and only if $\Phi K_{01} = cK_S$ for some $c \neq 0$. It follows that $\Phi_\mu^*(X)$ maximizes $D_\mu(\Phi)$ if and only if there exists $c \neq 0$ such that

$$\begin{aligned} cK_S(t,s) &= \Phi_\mu^* K_{01}(t,s) \\ &\triangleq \lim_{N \to \infty} \sum_{i,j=1}^{\infty} \phi_{ij}^*(N) K_0(t,t_i^*) K_1(s,t_j^*). \end{aligned}$$

As with Φ_α^*, it may be that no such Φ_μ^* exists, but we always have

$$D_\mu^* \triangleq \sup_{\Phi \in \mathbb{F}(K_{01})} D_\mu(\Phi) = \tfrac{1}{2} \langle K_S, K_S \rangle_{H(K_0) \otimes H(K_1)}.$$

Now, if Problem (D.1.1) is strongly nonsingular, then we know that there exists an optimal quadratic statistic $\Phi_{opt}(X)$ satisfying Equations (D.2.8) and (D.2.9). Since (D.2.9) clearly implies that $\Phi_{opt} K_{01} = K_S$ in $H(K_0) \otimes H(K_1)$, it follows that $\Phi_{opt}(X)$ maximizes $D_\mu(\Phi)$.

Conversely, it follows fairly easily from the results in [23] and [57] that, even if (D.1.1) is not strongly nonsingular, a quadratic statistic $\Phi(X)$ maximizes $D_\mu(\Phi)$ only if the corresponding detector is optimal.

For the signal detection problem, it is easy to verify that D_μ can be rewritten as

$$D_\mu(\Phi) = \frac{\left[E_1\{\Phi(X)\} - E_0\{\Phi(X)\}\right]^2}{½\mathrm{Var}_0\{\Phi(X)\} + ½\left[\mathrm{Var}_1\{\Phi(X)\} - \mathrm{Var}_S\{\Phi(X)\}\right]}, \tag{D.3.3}$$

where $\mathrm{Var}_S\{\Phi(X)\}$ is the variance of $\Phi(X)$ when only the signal is present. In this form, it is clear that D_μ is equivalent to the so-called *modified deflection ratio* (see [17]). The denominator in (D.3.3) can be interpreted as the average of the variance of $\Phi(X)$ in the presence of noise only and the variance of $\Phi(X)$ due to noise in the presence of the signal. Viewed in this fashion, D_μ becomes somewhat more intuitively appealing.

As a final remark, we note that, while it is not directly related to detector error probability in any obvious way, D_μ may be, in general, a better predictor of detector performance than the GSNR. In particular, for the case of slow and fast-fading channels considered by Gardner in [16], D_μ displays significantly less anomalous behavior than the GSNR when viewed as an indicator of performance gain with increasing observation time. In general, this conjecture is based on the fact that $D_\mu(\Phi)$ is essentially a measure of the distance, in $\boldsymbol{H}(K_0) \otimes \boldsymbol{H}(K_1)$, between ΦK_{01} (suitably normalized) and the nearest constant multiple of K_S. To be more precise,

$$d^2(\Phi K_{01}, K_S) \triangleq \inf_{c \in \boldsymbol{R}} \left\| \frac{\Phi K_{01}}{\|\Phi K_{01}\|} - cK_S \right\|^2$$

$$= 1 - \frac{2}{\|K_S\|^2} D_\mu(\Phi),$$

where $\|\cdot\|$ is the norm in $\boldsymbol{H}(K_0) \otimes \boldsymbol{H}(K_1)$. Hence, if Φ_1 and Φ_2 are two quadratic statistics and

$$D_\mu(\Phi_1) > D_\mu(\Phi_2),$$

then

$$d^2(\Phi_1 K_{01}, K_S) < d^2(\Phi_2 K_{01}, K_S).$$

Since the optimal test statistic is essentially defined in $\boldsymbol{H}(K_0) \otimes \boldsymbol{H}(K_1)$ and corresponds to K_S, the detector incorporating Φ_1 can be regarded as being "closer" to the optimal detector than the detector incorporating Φ_2. Heuristically then, D_μ can be regarded as a measure of the deviation of a quadratic detector from the optimal. On this basis, one might expect D_μ to be a fair indicator of detector performance. (See [11] for a discussion of a similar property of the *efficacy* of nonlinear detectors for deterministic signals.)

REFERENCES

[1] N. Aronszajn, "Theory of reproducing kernels," *Transactions of the American Mathematical Society*, vol. 68, pp. 337-404, 1950.

[2] R. B. Ash and M. F. Gardner, *Topics in Stochastic Processes*. New York: Associated Press, 1975.

[3] C. R. Baker, "Optimum quadratic detection of a random vector in Gaussian noise," *IEEE Transactions on Communication Technology*, vol. COM-14, pp. 802-805, 1966.

[4] V. Barbu and Th. Precupanu, *Convexity and Optimization in Banach Spaces*. Alphen Rijn, Holland: Sijthoff and Noordhoff, 1978.

[5] J. A. Barnes and D. W. Allan, "A statistical model of flicker noise," *Proceedings of the IEEE*, vol. 54, pp. 176-178, February 1966.

[6] R. J. Barton and H. V. Poor, "On generalized signal-to-noise ratios in quadratic detection," to appear.

[7] J. H. Beder, "A sieve estimator for the mean of a Gaussian process," *The Annals of Statistics*, vol. 15, pp. 59-78, 1987.

[8] J. Beran and H.-R. Kunsch, "Location estimators for processes with long-range dependence," Research Report no. 40, Swiss Federal Institute of Technology, Zurich, March 1985.

[9] C. Chen and S. A. Kassam, "Robust Wiener filtering for multiple inputs with channel distortion," *IEEE Transactions on Information Theory*, vol. IT-30, pp. 674-677, July 1984.

[10] C. Chen and S. A. Kassam, "Robust multiple-input matched filtering: Frequency and time-domain results," *IEEE Transactions on Information Theory*, vol. IT-31, pp. 812-821, November 1985.

[11] S. V. Czarnecki and K. S. Vastola, "Approximation of locally optimum detector nonlinearities," *IEEE Transactions on Information Theory*, vol. IT-31, pp. 835-838, November 1985.

[12] C. DeBoor and R. Lynch, "On splines and their minimum properties," *Journal of Mathematical Mechanics*, vol. 15, pp. 953-969, 1966.

[13] D. L. Donoho and R. C. Liu, "The automatic robustness of minimum distance functionals," *Annals of Statistics*, vol. 16, pp. 552-586, 1988.

[14] N. Dunford and J. T. Schwartz, *Linear Operators*. New York: Interscience Publishers, Inc., 1963.

[15] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*. New York: John Wiley & Sons, 1984.

[16] W. A. Gardner, "Anomalous behavior of receiver output SNR as a predictor of signal detection performance exemplified for quadratic receivers and incoherent fading Gaussian channels," *IEEE Transactions on Information Theory*, vol. IT-25, pp. 743-745, 1979.

[17] W. A. Gardner, "A unifying view of second-order measures of quality for signal classification," *IEEE Transactions on Communications*, vol. COM-28, pp. 807-816, 1980.

[18] H. Graf, "Long-Range Correlations and Estimation of the Self-Similarity Parameter," Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich, 1983.

[19] H. Graf, F. R. Hampel, and J.-D. Tacier, "The problem of unsuspected serial correlations," in *Robust and Nonlinear Time Series Analysis*, ed., J. Franke, W. Haerdle, D. Martin. New York: Springer-Verlag, pp. 127-145, 1984.

[20] U. Grenander, *Abstract Inference*. New York: John Wiley & Sons, 1981.

[21] J. Hajek, "On linear statistical problems in stochastic processes," *Czechoslovakian Mathematics Journal*, vol. 12, pp. 404-444, 1962.

[22] T. Kailath, "An RKHS approach to detection and estimation problems - Part I: deterministic signals in Gaussian noise," *IEEE Transactions on Information Theory*, vol. IT-17, pp. 530-549, 1971.

[23] T. Kailath and H. L. Weinert, "An RKHS approach to detection and estimation problems - Part II: Gaussian signal detection," *IEEE Transactions on Information Theory*, vol. IT-21, pp. 15-23, 1975.

[24] S. A. Kassam, T. L. Lim, and L. J. Cimini, "Two-dimensional filters for signal processing under modeling uncertainties," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-18, pp. 331-336, October 1980.

[25] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proceedings of the IEEE*, vol. 73, pp. 433-481, March 1985.

[26] M. S. Keshner, "Renewal Process and Diffusion Models of 1/f Noise," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, 1979.

[27] N. Kono, "Hausdorff dimension of sample paths for self-similar processes," in *Dependence in Probability and Statistics*, ed., E. Eberlein and M. S. Taqqu. Boston: Birkhauser, 1986.

[28] H.-R. Kunsch, "Statistical Aspects of Self-Similar Processes," Research Report no. 51, Swiss Federal Institute of Technology, Zurich, 1986.

[29] H.-R. Kunsch, "Discrimination between monotonic trends and long-range dependence," *Journal of Applied Probability*, vol. 23, pp. 1025-1030, 1986.

[30] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*. Cambridge, MA: MIT Press, 1983.

[31] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: John Wiley & Sons, 1969.

[32] T. Lundahl, W. Ohley, S. M. Kay, and R. Siffert, "Fractional Brownian motion: A maximum likelihood estimator and its application to image texture," *IEEE Transactions of Medical Imaging*, vol. MI-5, pp. 152-161, September 1986.

[33] B. B. Mandelbrot, "Self-similar error clusters in communication systems and the concept of conditional stationarity," *IEEE Transactions on Communication Technology*, vol. COM-13, pp. 71-90, March 1965.

[34] B. B. Mandelbrot, "Some noises with 1/f spectrum, a bridge between direct current and white noise," *IEEE Transactions on Information Theory*, vol. IT-13, pp. 289-298, April 1967.

[35] B. B. Mandelbrot, *The Fractal Geometry of Nature*. San Francisco: W. H. Freeman, 1983.

[36] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, vol. 10, pp. 422-437, October 1968.

[37] D. L. Mohr, "Modeling Data as a Fractional Gaussian Noise," Ph.D. dissertation, Princeton University, Princeton, 1981.

[38] G. M. Molcan and Ju. I. Golosov, "Gaussian stationary processes with asymptotic power spectrum," *Soviet Mathematics Doklady*, vol. 10, pp. 134-137, 1969.

[39] G. V. Moustakides and S. A. Kassam, "Robust Wiener filters for random signals in correlated noise," *IEEE Transactions on Information Theory*, vol. IT-29, pp. 614-619, July 1983.

[40] G. V. Moustakides and S. A. Kassam, "Minimax equalization for random signals," *IEEE Transactions on Communications*, vol. COM-33, pp. 820-825, August 1985.

[41] A. W. Naylor and G. R. Sell, *Linear Operator Theory in Engineering and Science*. New York: Springer-Verlag, 1982.

[42] G. L. O'brien and W. Vervaat, "Marginal distributions of self-similar processes with stationary increments," *Zeitschrift fur Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 64, pp. 129-138, 1983.

[43] G. L. O'brien and W. Vervaat, "Self-similar processes with stationary increments generated by point processes," *The Annals of Probability*, vol. 13, pp. 28-52, 1985.

[44] K. B. Oldham and J. Spanier, *The Fractional Calculus*. New York: Academic Press, 1974.

[45] H. Oodaira, "The Equivalence of Gaussian Stochastic Processes," Ph.D. dissertation, Michigan State University, East Lansing, 1963.

[46] A. Papamarcou and T. L. Fine, "A note on undominated lower probabilities," *The Annals of Probability*, vol. 14, pp. 710-723, 1986.

[47] E. Parzen, *Time Series Analysis Papers*. San Francisco: Holden-Day, 1967.

[48] E. Parzen, "Statistical inference on time series by RKHS methods," in *Proceedings of the 12th Biennial Seminar of the Canadian Mathematical Conference*, Montreal, pp. 1-37, 1970.

[49] S. Peleg, J. Naor, R. Hartley, and D. Avnir, "Multiple resolution texture analysis and classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, pp. 518-523, July 1984.

[50] A. P. Pentland, "Fractal-based description of natural scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, pp. 661-674, November 1984.

[51] B. Picinbono and P. Duvaut, "Optimal linear-quadratic systems for detection and estimation," *IEEE Transactions on Information Theory*, vol. 34, pp. 304-311, March 1988.

[52] H. V. Poor, "On robust Wiener filtering," *IEEE Transactions on Automatic Control*, vol. AC-25, pp. 531-536, June 1980.

[53] H. V. Poor, "Robust matched filters," *IEEE Transactions on Information Theory*, vol. IT-29, pp. 677-687, September 1983.

[54] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1988.

[55] M. Rosenblatt, "Independence and dependence," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. II, Berkeley, pp. 431-443, 1961.

[56] Yu. A. Rozanov, "On probability measures in functional spaces corresponding to stationary Gaussian processes," *Theory of Probability and its Applications*, vol. IX, pp. 405-420, 1964.

[57] Yu. A. Rozanov, *Infinite-Dimensional Gaussian Distributions*. Providence, RI: American Mathematical Society, 1971.

[58] L. A. Shepp, "Radon-Nikodym derivatives of Gaussian measures," *Annals of Mathematical Statistics*, vol. 37, pp. 321-354, 1966.

[59] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*. Princeton, NJ: Princeton University Press, 1970.

[60] M. S. Taqqu, "A representation for self-similar processes," *Stochastic Processes and their Applications*, vol. 7, pp. 55-64, 1978.

[61] M. S. Taqqu, "Self-similar processes and related ultraviolet and infrared catastrophes," *Random Fields: Rigorous Results in Statistical Mechanics and Quantum Field Theory*, vol. 27, pp. 1057-1096, 1981.

[62] M. S. Taqqu, "Self-Similar Processes," in *Encyclopedia of Statistical Sciences*. 1985.

[63] M. S. Taqqu, "A bibliographical guide to self-similar processes and long-range dependence," in *Dependence in Probability and Statistics*, ed., E. Eberlein and M. S. Taqqu. Boston: Birkhauser, 1986.

[64] M. S. Taqqu and C. Czado, "A survey of functional laws of the iterated logarithm for self-similar processes," *Stochastic Models*, vol. 1, pp. 77-115, 1985.

[65] K. S. Vastola, "On robust Wiener signal estimation," *IEEE Transactions on Automatic Control*, vol. AC-31, pp. 466-467, May 1986.

[66] K. S. Vastola and H. V. Poor, "Robust Wiener-Kolmogorov theory," *IEEE Transactions on Information Theory*, vol. IT-30, pp. 316-327, March 1984.

[67] S. Verdu, "Comments on 'Anomalous behavior of receiver output SNR as a predictor of signal detection performance exemplified for quadratic receivers and incoherent fading Gaussian channels'," *IEEE Transactions on Information Theory*, vol. IT-28, pp. 952-953, November 1982.

[68] S. Verdu and H. V. Poor, "Minimax robust discrete-time matched filters," *IEEE Transactions on Communications*, vol. COM-31, pp. 208-215, February 1983.

[69] S. Verdu and H. V. Poor, "Signal selection for robust matched filtering," *IEEE Transactions on Communications*, vol. COM-31, pp. 667-670, May 1983.

[70] S. Verdu and H. V. Poor, "On minimax robustness: A general approach and applications," *IEEE Transactions on Information Theory*, vol. IT-30, pp. 328-340, March 1984.

[71] W. Vervaat, "Sample path properties of self-similar processes with stationary increments," *The Annals of Probability*, vol. 13, pp. 1-27, 1985.

[72] R. F. Voss and J. Clarke, "1/f noise in music and speech," *Nature*, vol. 258, pp. 317-318, November 1975.

[73] C. L. Weber, *Elements of Detection and Signal Design*. New York: McGraw-Hill, 1968.

[74] H. L. Weinert, Ed., *Reproducing Kernel Hilbert Spaces: Applications in Statistical Signal Processing*. New York: Nostrand Reinhold Company, 1982.

[75] W. E. Williams, "A class of integral equations," *Proceedings of the Cambridge Philosophical Society*, vol. 59, pp. 589-597, 1963.

[76] D. Wolf, Ed., *Noise in Physical Systems*. New York: Springer-Verlag, 1978, Proceedings of the Fifth International Conference on Noise, Bad Nauheim, March 1978.

[77] A. M. Yaglom, "Correlation theory of processes with stationary random increments of order n," *American Mathematical Society Translations, Series 2*, vol. 8, p. 87, 1958.

[78] A. M. Yaglom, *An Introduction to the Theory of Stationary Random Functions*. Englewood Cliffs, NJ: Prentice-Hall, 1962.

[79] K. Yao, "Applications of reproducing kernel Hilbert spaces - Bandlimited signal models," *Information and Control*, vol. 11, pp. 429-444, 1967.